

A view on global urbanization. Spatial networks, urban scaling and remote sensing.

THÈSE N° 7067 (2016)

PRÉSENTÉE LE 23 NOVEMBRE 2016

À LA FACULTÉ DE L'ENVIRONNEMENT NATUREL, ARCHITECTURAL ET CONSTRUIT
LABORATOIRE DE SYSTÈMES D'INFORMATION GÉOGRAPHIQUE
PROGRAMME DOCTORAL EN GÉNIE CIVIL ET ENVIRONNEMENT

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Emanuele STRANO

acceptée sur proposition du jury:

Prof. N. Geroliminis, président du jury
Prof. F. Golay, Prof. A. Rinaldo, directeurs de thèse
Prof. M. Batty, rapporteur
Prof. V. Latora, rapporteur
Prof. P. De Los Rios, rapporteur



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2016

"... and every new fact,
illustrative of the action and reaction between humanity and the material world around it,
is another step toward the determination of the great question,
whether man is of nature or above her."
G.P. Marsh, 1864

To Anna.

Acknowledgements

To my supervisors Prof. Francois Golay and Prof. Andrea Rinaldo an immense thanks. I also truly thanks all the Professors that at various stages of this work helped me: Sergio Porta, Vito Latora, Marc Barthelemy, Enrico Bertuzzo, Josè Lobo, Mike Batty, Karen Seto, Paolo De Los Rios, Marta Gonzales, Elisabeth Wentz, Andy Adamatzky and Andrew Tatem. Young fellows have been of great help and without them I wouldn't have moved a single step. I had the privilege to work together with some of the most promising scientists of the next generation: Sarai Shay, Matheus Viana, Andrea Giometto, Silvia Zaoli, Matheus Parkan, Thomas Esch, Alessio Cardillo, Davide Cucci, Leonardo Bellocchi, Bilal Farooq, and Vishal Sood. I thanks also Veronique for her constant and preciod help.

Lausanne, April 2016

Emanuele Strano

Abstract

It is widely recognized that global urbanization is one the most urgent issue of our time. Any problem related to global issues, such as climate change, social and environmental sustainability and the future of our society, is somehow related to cities and urbanization. However, despite its great importance, general knowledge on cities still lacks a scientific and shared understanding. Traditional approach for understanding cities, mostly derived from urban planning practices and classical geography, have been proved to be unsuited to catch the great complexity of cities. In order to fill this gap, and thanks to the increasing amount of digital data, a new science of cities is arising. Embedded in physic and chemistry approaches, the new science of cities aims to provide quantitative and reproducible knowledge on urban form and dynamics. From urban scaling to complex networks, scientific approach to cities was able to provide evidence of the very top-down nature of urbanization and of the existence of some physical laws that govern urbanization beyond local conditions. The thesis, acknowledging the weak points of traditional approaches to cities, aims to provide a series of quantitative studies on urbanization patterns and urban dynamics. Six case studies explore several open questions providing new insights on the physical nature of urbanization process. The first case study provides an empirical analysis of the evolution of a road network over 200 years. It is shown that few processes are regulating urban roads evolution which are independent of technological transport evolution and population growth. The second case studies introduces a new metrics for planar networks which has been used to compare artificial and natural transportation networks. It has been shown that, despite the spatial constraints given by planarity, human made networks are similar each others and they are more navigable than other natural networks. The third case study provides the analysis of the multiplex networks composed by roads plus underground in the metropolitan areas of London and New York. The study provides evidence that the multiplex nature of urban networks, despite the fact that is has been rarely taken in account, strongly effects the displacements of urban centrality thus being of great importance for any realistic urban grow model. The fourth case study reports a time analysis of the scaling relationship between population and gross metropolitan product (GMP) in 240 European cities. It has been shown that population/GMP scaling is not universally super-linear, as claimed by several recent publications. Time analysis of scaling relationship shows that in the emerging economies of East Europe scale relationship is super-linear and increase of time. For stable economies instead, as for West European cities, population/GMP is linear and stable over time. The dualistic scaling behavior is not mirrored by the spatial organization of town. This in turn suggests that the economic performance of cities is independent of urban form. The fifth case study is the structural analysis of the network made by all major roads on Earth. Coupling a global and updated inventory of road network with global inventory of land-use, it is shown that the global road network has a self-similar shape among land-uses and location. The self similar nature of this huge infrastructure is due to processes of hierarchical

Acknowledgements

fragmentation. The study provides evidence of the universality of such fragmentation process which in turn confirm that road evolution is governed by simple and universal laws. Finally, in the sixth case study, the scaling analysis of all settlements on Earth is provided. Preliminary results show the self similar nature of global urbanization, however, a regional analysis shows that in highly urbanized regions self-similar functional form are not valid. Results suggests that power law distribution of urban settlements is not an universal feature and that in some regions, specially in China, urban growth might have reached its tip point without displaying any maximum size. Perspectives for further analyses close the thesis, in particular the hypothesis of the global trade network as determinate of global urbanization is proposed.

Key words: Global urbanization, urban form, complex spatial networks, urban scaling.

Contents

Acknowledgements	i
Abstract	iii
List of figures	vii
List of tables	ix
Introduction: Cities and global urbanization	1
1 State of art and open problems	7
1.1 Remote sensing and inventories of global urbanization	7
1.2 Urban scaling: Zipf's law, allometry, and the importance of being large	14
1.3 Complex spatial networks and urban form	20
1.4 Main gaps and open problems	28
2 Material and methods	29
2.1 Road and transportation networks	29
2.2 Land-use data	34
3 Results	39
3.1 Evolution of roads network, an empirical analysis	40
3.1.1 Introduction	40
3.1.2 Results	41
3.1.3 Discussion	49
3.2 The simplicity of planar networks.	52
3.2.1 Introduction	52
3.2.2 Material and methods	53
3.2.3 Results	56
3.2.4 Discussion	61
3.3 Multiplex approach to urban network analysis	63
3.3.1 Introduction	63
3.3.2 Material and Methods	64
3.3.3 Results	65
3.3.4 Discussion	75

Contents

3.4	Scaling of urban economic performance over time and its relationship with the urban land-use	76
3.4.1	Introduction	76
3.4.2	Material and methods	77
3.4.3	Results	77
3.4.4	Discussion	82
3.5	Structural proprieties and scaling of the global road network	85
3.5.1	Introduction	85
3.5.2	Materials and methods	86
3.5.3	Results	87
3.5.4	Discussion	92
3.6	Testing the universality of urban scaling.	95
3.6.1	Introduction	95
3.6.2	Results	96
3.6.3	Discussion	102
3.6.4	Perspective: A new inventory of global human settlements	102
4	Conclusions and perspectives	105
4.1	Conclusions: Towards the development of a science for cities	105
4.2	Perspectives: The hypothesis of the global trade network as global urbanization determinant	107
	Bibliography	118
	Curriculum Vitae	119

List of Figures

1	Global urban transition	1
1.1	An inventory of first generation global urban footprint as reported by Potere et al.	12
1.2	Comparison between first and second generation global urban inventories products. . .	13
1.3	Typical representation of Zipf's law	15
1.4	Modelling Zipf's law	16
1.5	Sub and super linear scaling for cities	18
1.6	Primal and dual urban street patterns	21
1.7	Cost and efficiency describe historical road patterns.	22
1.8	Roads' length distribution of ten European cities	23
1.9	Centralities in a spatial network	25
1.10	Centralities vs. commerce in Bologna	27
2.1	Historical map sequences for Groane	31
2.2	Global road network map	33
2.3	Biological network used in the study.	33
2.4	Location and map of the 300 European cities	34
2.5	Global Land Cover Dataset	36
2.6	GL30 preparation	37
2.7	Global crop-land spatial inventory.	38
3.1	Outline of the case studies	39
3.2	Evolution of road network in the Groane area (Milan, Italy)	42
3.3	Network evolution measures	43
3.4	Evolution of road cells	45
3.5	Cumulative distributions of length evolution	47
3.6	Densification and exploration	48
3.7	Betweenness centrality vs. road's age	50
3.8	Simplest path: example and calculation	54
3.9	Dual vs primal network degree distribution	56
3.10	Illustration of the null model of a simplicity value	57
3.11	Simplicity profiles	59
3.12	Simplicity profiles for time-varying networks	61
3.13	Simplicity index	62
3.14	The spatial extent of the two metropolitan areas	64

List of Figures

3.15 Average interdependence	68
3.16 Average local outreach	69
3.17 Spatial distribution of the local outreach in London	70
3.18 Local outreach	71
3.19 The spatial distribution of BC	74
3.20 Land-use example cities case	78
3.22 Location of the 248 European cities case	80
3.21 Linear vs super linear scaling	81
3.23 Scaling relationships between main urban land uses and the total urban area	83
3.24 Global road network data preparation	88
3.25 Evidence for a universal length distribution.	90
3.26 Scaling by macro regions	91
3.27 Topological road structure for all countries.	92
3.28 Hierarchical fragmentation of the global road network.	93
3.29 Scaling of all settlements on Earth	98
3.30 Scaling for low and high resolution data.	99
3.31 Urban vs. rural patterns in Niger and Nigeria	101
3.32 A preview of GUF at global scale	103

List of Tables

1.1	Global urban inventories	11
2.1	List of geographical information sources used to construct the Groane data set.	30
3.1	Basic features of multiplex urban network	65
3.2	Scaling for low and high resolution urban patches	100

Introduction

One of the iconic images that best describes human conditions in our time is a simple graph composed of two lines representing the urban and rural percentages of the global population between 1950 and 2050. In 1950, the global rural population was double the urban population. After that time, it has been constantly decreasing, with a consequent increase in the urban population. By 2010, these two lines had crossed each other, and the urban population began to exceed the rural. This observation, commonly known as the global urban transition [1], indicates that the majority of people on Earth inhabit some kind of urban environment.

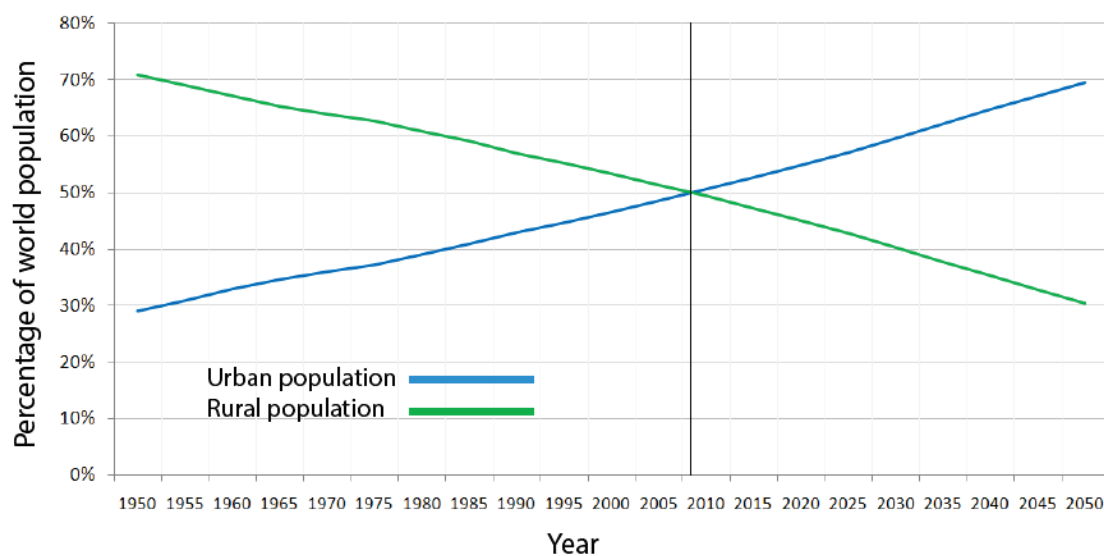


Figure 1: **The global urban transition.** One of the iconic and widely used representation of the global urban population transition. Image by Taylor Luker via Wikimedia Commons, data from United Nation Population Division [2].

Urbanization is an ancient phenomenon, and humans have been building cities for the last three thousand years. In developed countries such as the US, the predominance of cities can be tracked back to the earliest decades of the 19th century, whereas in Britain, this predominance dates to the middle of the 18th century [3]. However, as summarized by Seto et al. [4], the contemporary global process of urbanization exhibits certain special characteristics that differentiate it from that of the past: i) The scale of urbanization: Cities are much larger than in any previous time. ii) The growth

rate: cities are growing much more rapidly than before. iii) The locations of urbanization: Cities are developing most rapidly and extensively in Africa and Asia. Arguments regarding the magnitude of global urbanization [5, 6, 7, 8] have made clear to the global scientific and policy-making community that cities play a primary role as drivers of all social and environmental systems. Currently, there is considerable evidence that global urbanization impacts the entire spectrum of human and natural systems. By increasing demands for energy, water, and food, cities directly affect most ecosystems [9, 10] at the local, meso and global scales [11], with global impacts on biodiversity, climate [12, 13] and human health [14]. Cities are also economic drivers. It has been argued that up to 90% of global gross added value comes from urban areas [15]. Population projections make such figures even more interesting and alarming. It has been proposed that by 2050, the global population will grow to 9 billion [2, 1], and considering the urban/rural transition, studies have inferred that the majority of this additional population will live in cities [16, 17]. All of these considerations place the global urbanization process in a crucial position regarding global sustainability issues. Beyond the *global urbanization issue* mantra, which is largely used to promote research agendas focusing on cities, a fundamental question pertains to knowledge of the global extent of urbanization as well as its dynamics and drivers. It is clear that any possible strategies for controlling and minimizing the effects of the massive upcoming urbanization must depend on the predictive capabilities regarding it. Many scholars, however, share the view that estimations of the effects of human presence on Earth are strongly biased [18, 19] and that the dynamics of growth [20] and its economic and social effects [21] are poorly understood. In fact, even a brief review of the major literature reveals uncertainties in assessment. For example, the iconic picture of the global urban transition is based on urban-rural censuses of the global population. These data have been stored by several international organizations, such as the World Bank, [22] but only the United Nations Population Division (UNPD) provides the urban/rural population ratio [2]. The sources of data for the UNPD are the latest available official censuses for each country. Because the UNPD data are the only available data of this kind, they have been extensively used to estimate and forecast global urbanization [23, 17, 24]. However, these data have also been subject to several criticisms, as reported by Fragkias et al. [25]. The most relevant criticism, as the UN itself recently admitted [1], is the lack of a consistent definition of urban and rural populations shared across all countries. For example, in a 228-country survey of the UN rural/urban statistics, more than 10 definition of urban population were identified [26]. This lack of consistency, in turn, produces systematic biases regarding the actual state of urbanization and, in a cascade effect, on analyses of urbanization drivers. In other words, the use of non-standardized census surveys can potentially produce biased results. Moreover, as suggested by Montgomery et al. [16] and demonstrated by Tatem et al. [27], biases with respect to urban population censuses are particularly prone to arise in developing countries, where censuses are affected by informal urbanization (e.g., slums) and the high rate of population growth.

A second source of uncertainty in the state of global urbanization is the use of spatially explicit methods to analyse urban land use at the global scale. Urban land-use surveys provide the possibility of measuring the forms that cities take, their spatial evolution and the ways in which they expand to fill up free and open land. In spatially explicit approaches, optical imagery is predominantly used to detect urban areas and couple these areas with demographic projections. These approaches suffer from a double bias: the uncertainties of imagery-based land cover classification [18, 19] and the census biases described above. These biases clearly arise when comparing the findings of important studies on the global extent of urban land cover and projections. Seto et al. [23], using a probabilistic approach,

argued that the global extent of urbanization will increase over the next twenty years by 85% to 155%. Another important analysis, performed by Angel et al. [17] using regression methods, suggests that over the same period, global urban land cover may increase by 55% to 183%. In a third study, Seto et al. reported that this percentage may be between 210% and 310% [24]. According to these results, one can conclude that over the next twenty years (a very short time period), global urbanization cover may increase by anywhere from 55% to 310%!

A third aspect of the debate on global urbanization, which is very important but appears not to have yet been fully clarified, concerns the economic performance and drivers of cities. The established fact that the urbanization in developed countries was accompanied by economic growth and industrialization has generated an expectation of a universal virtuous circle between economic growth (measured mainly in terms of gross domestic product) and urbanization, regardless of local conditions [28, 29]. From classical urban economic theories [30, 31] to the more recent scaling approach to cities [20], the growth of the urban population has routinely been used as a proxy for economic growth and vice versa. However, this pattern cannot be regarded as universal [32], and the deviations from it have not yet been fully explained. In fact, as recently noted in several studies [33, 16, 34, 35], the increasing rates of urbanization in persistently poor and non-industrialized countries pose an important dilemma regarding classical urban economic theories and paradigms. For example, why are the rates of economic growth in Kinshasa and Dhaka, both of which have a population of more than 10 million, nearly stagnant? Why, despite their similar rates of urbanization, is Asia host to a number of explosive economies while sub-Saharan Africa has seen very little economic growth? These patterns have not been explained thus far. These three examples make it clear that the overall figures on the extent of urbanization and its drivers are highly uncertain and warrant further investigations. Moreover, the general debate on global urbanization is lacking an objective and sufficiently scientific view.

Upon returning to the three points that differentiate contemporary urbanization from the urbanization of the past, several questions arise: i) *Cities are much larger than in any previous time*. This statement is true, but it implies the assumption that cities, at some stage of development, should cease growing in size. No scientific justification or motivation for this assumption has been found thus far. ii) *Cities are growing much faster than before*. This is also true, but it does not imply that cities are following a different growth dynamic than they have in the past. Any population with a fixed rate of growth that is independent of the size of the settlement will grow exponentially, meaning that over time, it will grow increasingly rapidly; this means that the way in which cities are currently growing may be fully consistent with the way in which they have grown in the past, and further analysis is needed to claim that contemporary urbanization processes are truly different. *Cities are spreading most rapidly in Africa and India*. Why should this not be the case? What precludes China, India or Africa from undergoing a similar process of urbanization as in Europe or the US?

By acknowledging the fundamental relevance of a deeper understanding of urbanization patterns and moving forward from the mere observation of controversial and unclear knowledge of these patterns, the thesis aims to contribute to the development of a better understanding of the form and the drivers of urbanization patterns and cities. The thesis is particularly focused on the spatial aspects of urbanization, with demographic and social data regarded as ancillary information. The research can be framed in terms of three major, broad areas of research: urban remote sensing, complex spatial networks and urban scaling.

The thesis is organized in four main chapters as described below:

- **Chapter 1** provides the state of the art of the three main areas of research including the main gaps and open questions. Chapter 1 includes also specific research questions as well as specific research hypotheses and motivations.
- In **Chapter 2** a full overview of materials and methods is provided. This section reports the source of data and clarify their different nature. It also provides specific methodological aspects to prepare and fuse them.
- **Chapter 3** presents the results and it is organized in several subsections. Each of them reports a specific and self-contained case study. Each case study is simply organized with an introduction followed by methods, results and specific discussion. The natural order of presentation follows the time in which results has been produced that couples a spatial scale order. Results appearance moves from local scale analyses, to regional and global scale. The following case studies are reported:
 - **An empirical analysis of the evolution of road network over 200 years.** In this study an evolving road network has been extracted from historical cartography and topographic maps. 200 years of evolution has been mapped over 6 time steps. Empirical data on road network evolution allowed to analyze its the structural proprieties over time that are ultimately correlated to population growth and urbanization processes. The study demonstrates that urbanization, land fragmentation are parts of the same process of the road network evolution. Moreover roads' evolution processes seems to be invariant over the entire period, thus independent of transportation technology. s
 - **The simplicity of planar networks and the comparison between man-made and natural transportation networks.** In this case a new metric for analyzing planar networks is introduced. This metric called *simplicity* measures the relative difference between length of simplest paths and shortest paths in planar graphs. Simplicity metric allows to produce a descriptive profile of planar networks that has been used to study differences between man-made and natural planar graphs. Road, rails and supply systems results much different than other natural biological transportation networks such as leaf venation. Results suggests that man-made systems optimize navigability.
 - **A multiplex approach to urban road networks.** Multiplex approach on road networks analysis is a relatively new problem in complex networks analysis. Similar to classical transportation analysis approach, the effect of underground system on road networks topology and centrality has been analyzed. Analyses have been performed on the New York and London metropolitan areas. Results shows a great change in the geography of urban centrality, in the displacement of congested areas and the overall accessibility of cities. This case study demonstrates that multiplex transportation networks effects the entire network geography of the city suggesting an entire revision of precedent results.
 - **Time analysis of gross metropolitan product vs population and its correlation with urban land-use.** In this section the scaling relationship of gross metropolitan product vs

population has been explored over a six years period and for 300 European cities. Results shows a very interesting figure in which different scaling regimes overlap the West and East European geographical division. Moreover the study shows that super-linear scaling is not an universal urban feature posing important question on the universality of urban scaling. Scaling of GMP/pop also overlaps with road surface development of cities.

- **Structural proprieties of the global road network.** In this section the network composed by all major road on Earth is studies. This network can be seen as a good approximation of global urbanization. Results show a a self-similar structure that is independent of location and land-use.
- **Testing the universality of Zipf's law at global scale.** A global sample of cities has been extracted from remotely sensed data with the aim to test the universality of the probability distribution of cities' sizes. Results show that Zipf' s law is not universal and many regions on Earth follows different distribution. The study presents also a new global and high resolution data set of human settlements.
- **Chapter 4** is dedicated to extract and summarize research findings in a comprehensive framework. A set of conclusion, followed by perspectives for future works close the thesis.

1 State of art and open problems

1.1 Remote sensing and inventories of global urbanization

At present, urban remote sensing (URS) represents the most powerful tool available for obtaining precise spatial information on human settlements and land use patterns, especially in regions where no official maps are available and where cities grow rapidly and informally. URS encompasses the entire process of acquiring and analysing remotely sensed data in an urban context [36]. Remotely sensed data may be optical in nature, derived from satellite imagery, such as the data provided by Landsat satellites, but such data can also be obtained from radar sensors, such as Synthetic Aperture Radar (SAR) or laser scanning (LIDAR) instruments. The range of domains that fall under the umbrella of URS is wide, spanning from satellite engineering to image processing, from machine learning problems to GPS and orientation systems. It is outside the scope of this thesis to conduct a full review of URS, which would require a full thesis in itself. Instead, focus is placed on the specific aspects of URS related to measurements of urban areas at regional and global scales.

The global classification of urban settlements is a very specific topic in the realm of URS because of the trade-off between the resolution and the size of the study area. URS can operate over a wide range of resolutions (from 1 m to 1 km). Over the past two decades, the choice of resolution has been primarily application-dependent; for example, for any urban planning application, which is typically limited to a particular area of a city, high-resolution imagery (1 to 5 m) is both required and feasible. However, for regional and global analyses, a coarser-grained classification (500 m to 1 km) has been the only choice. Although representing an important source of information, given their low resolution coupled with the existence of different definitions of urban areas, the first generation of global urban mapping products suffered from inaccuracy and considerable disagreements. The sole complete comparative study of first-generation global urbanization mapping inventories to date was performed by Potere et al. [19]. Table 1.1 provides a list of the first generation of global urban mapping products, following Potere et al.[19], plus two more recent projects of the second, HR generation that are predominantly considered in this thesis. Below, several specifics are provided regarding each global urban product.

- **VMAPO** Vector Map Level Zero is a vectorial thematic data layer covering the entire surface of the Earth. Its features and attributes include major roads, railroads, hydrology systems,

utility networks, airports, coastlines, international boundaries and populated locations (cities). VMAP0 was developed by the National Imagery and Mapping Agency (NIMA) for the US military and is publicly available in its lower-resolution version. Its inventory of global urbanization is not sufficiently detailed for any analysis but rather provides a general idea of the locations of major cities.

- **GLC00** [37] Global Land Cover 2000 was produced under the auspices of the Fifth Framework Programme (1999-2002) for Research of the European Commission. It was developed by more than 30 collaborating groups using a bottom-up approach. The goal was to perform validation tests based on local regional expertise. It contains a set of raster thematic maps with 22 classes based on the classification of the SPOT-4 VEGETATION VEGA dataset provided by CNES (France). It has been documented that other remote data have been used to identify and locate large urban areas. For example, in Africa, Night Light sensors has been used.
- **LandScan** [38] The Land Scan Global Population Database is, at present, the only global inventory of urban surfaces that contains a proxy for population. It is a binary data raster at a resolution of approximately 1 km. The LandScan project is continuously evolving, and it incorporates many sources of spatial information; interpolation algorithms are constantly being developed to classify the extent of urbanization and to assign population values on the basis of proximity to transportation networks, water systems, slopes and other factors. The information layers are of three kinds: vectorial layers, typically containing political boundaries and coastlines; imagery layers, obtained from instruments ranging from radar sensors to IKONOS; and tables containing data such as census data and city population data. At present, LandScan is the most widely used global inventory of the human population; it has also been used to produce other refined global products, such as HYDE3 and IMPSA. LandScan was maintained and updated from 1998 to 2012. Despite its widespread use, its low resolution precludes its application to a precise analysis of rural and low-density settlements.
- **LITES** [39] The LITES product offers a continuous layer representing the intensity of stable night light provided by the Earth Observation Group (EOG) at the National Oceanic and Atmospheric Administration (NOAA), US. It is available at the global level with a 1 km resolution. Stable night light is usually associated with urban land cover. LITES include a time inventory of global urban settlements from 1994 to the present time. LITES is produced by means of semi-automatic cloud detection and filtering of Operational Linescan System (OLS) satellite images; OLS imagery covers the visible and infrared regions of the spectrum. LITES has been extensively used to measure global urbanization; however, its low resolution poses a considerable problem for the recognition of small- or moderate-scale urban settlements. Another problem is related to the detection of non-illuminated urban settlements, such as slums and small towns in rural areas.
- **GlobCover**[40] is an initiative of the European Space Agency (ESA). It aims to provide a 300 m resolution map with 22 thematic classes, classified via automatic classification processes based on data obtained from the MERIS sensor onboard the ENVISAT environmental satellite (ESA). For many classes, it represents an improvement over the GLC00 map, but it has been documented that with regard to the urban class, GlobCover overlaps the GLC00

map without any improvement.

- **HYDE3** [41] The History Database of the Global Environment is an initiative developed under the direction of the Netherlands Environmental Assessment Agency. It is a continuous layer derived from LandScan data plus administrative boundaries that contain populations of between 1700 and 2000. The focus on time series and the use of political boundaries is most likely responsible for the production of a rather low-quality spatial product; in fact, HYDE3 cannot be used to track human footprints, only for large-scale models of changes in land use. It is available at a 10 km resolution.
- **IMPSA** The Global Impervious Surface Area product combines data from night light sensors with global population data based on LandScan to model the global extent of impervious surface areas. It was produced by the US National Geophysical Data Center.
- **GRUMP** The Global Urban Rural Mapping Project provides a binary map of urban-rural classes at 1 km resolution. It was developed by the Center for International Earth Science Information Network (CIESIN) at Columbia University using two data sources: 1) a human settlement database of approximately 55000 points of settlement with an estimated population of over 1000, for which the populations were obtained from official sources, and 2) a layer of 21000 urbanized zones extracted from night-time light data from 1994-1995.

As reported by Potere et al. and partially confirmed by this thesis, given the great heterogeneity in the classification of urban land among different products, there is a significant gap separating the different spatial information provided by the different products. Even in a visual analysis, large differences emerge, as can be seen in Fig.1.1. For example, the results of a simple estimation of the total urban extent vary by an order of magnitude between VMAP0 and GRUMP, from 0.3 to 3.4 million square kilometres. Potere et al. reported three main factors that cause such large inter-map differences: i) differences in the timing of map construction, ii) differences in map resolution and iii) large differences in the definition of urban land use. The disagreements between different definitions of urban land cover remain an open problem facing global urbanization studies; these disagreements arise from the spectral heterogeneity of urban environments [42] coupled with difficulties in defining urban boundaries. However, over the past few years, by virtue of the continually increasing availability and accessibility of high- or very-high-resolution remotely sensed data coupled with increasing computational power, a second generation of global and HR urban inventories has been developed by a few international agencies. To date, the best solution for global urbanization classification, at least for urban land cover classification, is provided by the German Aerospace Agency (DLR) in the form of the Global Urban Footprint dataset. GUF is a global binary settlement mask that indicates just-built and un-built-up areas at the very high spatial resolution of 0.4 arc-sec (12 m) and potentially at the global scale [43]. The classification method used to identify built-up areas makes use of the latest radar satellite sensors, called TerraSAR-X and TanDEM-X. Feature extraction is performed based on an analysis of local image heterogeneity [44]. The definition of urban land use is a non-trivial problem and is also a subject of this thesis. As explained in the Methods section, a different approach to defining urban land use can be provided by the use of GUF data, to overcome problems of spectral heterogeneity, coupled with vectorial data representing transportation and road networks. To conclude, a literature review of global urbanization inventories reveals that despite the considerable efforts

Chapter 1. State of art and open problems

of many international agencies, at present, a precise and reliable inventory of human presence on Earth is lacking.

1.1. Remote sensing and inventories of global urbanization

Table 1.1: The table reports all the global urban available inventories as reported by Potere et al, plus two more recent and up to dated global inventories.

Code	Name	First generation Urban Global Inventories	Owner/Producer	Type
VMAP0	Vector Map Level Zero	US National Intelligence	Geospatial-Intelligence	Thematic
GLC00	Global Land Cover 2000 v1.1	European Commission search	Join Re-	Thematic (22 classes)
GLOBC	GlobCover v2	European Commission search	Join Re-	Thematic (22 classes)
HYDE3	History data base on global env.	Netherlands Env. Assessment Agency		continuous
IMPSA	Global Impervious Surface Area	US National Geo-physic Data Center		continuous
MOD500	Modis Urban Land cover 500m	Uni. of Wisconsin, Boston Uni., NASA		binary map urban-rural
MOD1k	Modis Urban Land cover 1km	Boston Uni., NASA		binary map urban-rural
GRUMP	Global Rural-Urban Mapping Project	Earth institute at Columbia University		binary map urban-rural
LITES	Nighth-time Lights	National Gophysical Data Center, US		Continuous
LSCAN	LandScan 2005	US Oak Ridge National Laboratory		Continuous
Second generation				
GUF	Global Urban Footprints	DLR, German Aerospace Agency		Built-up classification, binary raster, 4 to 12m res.
GL30	Global Land 30m	National Geomatic Center of China		Thematic (9 classes), 30m res.

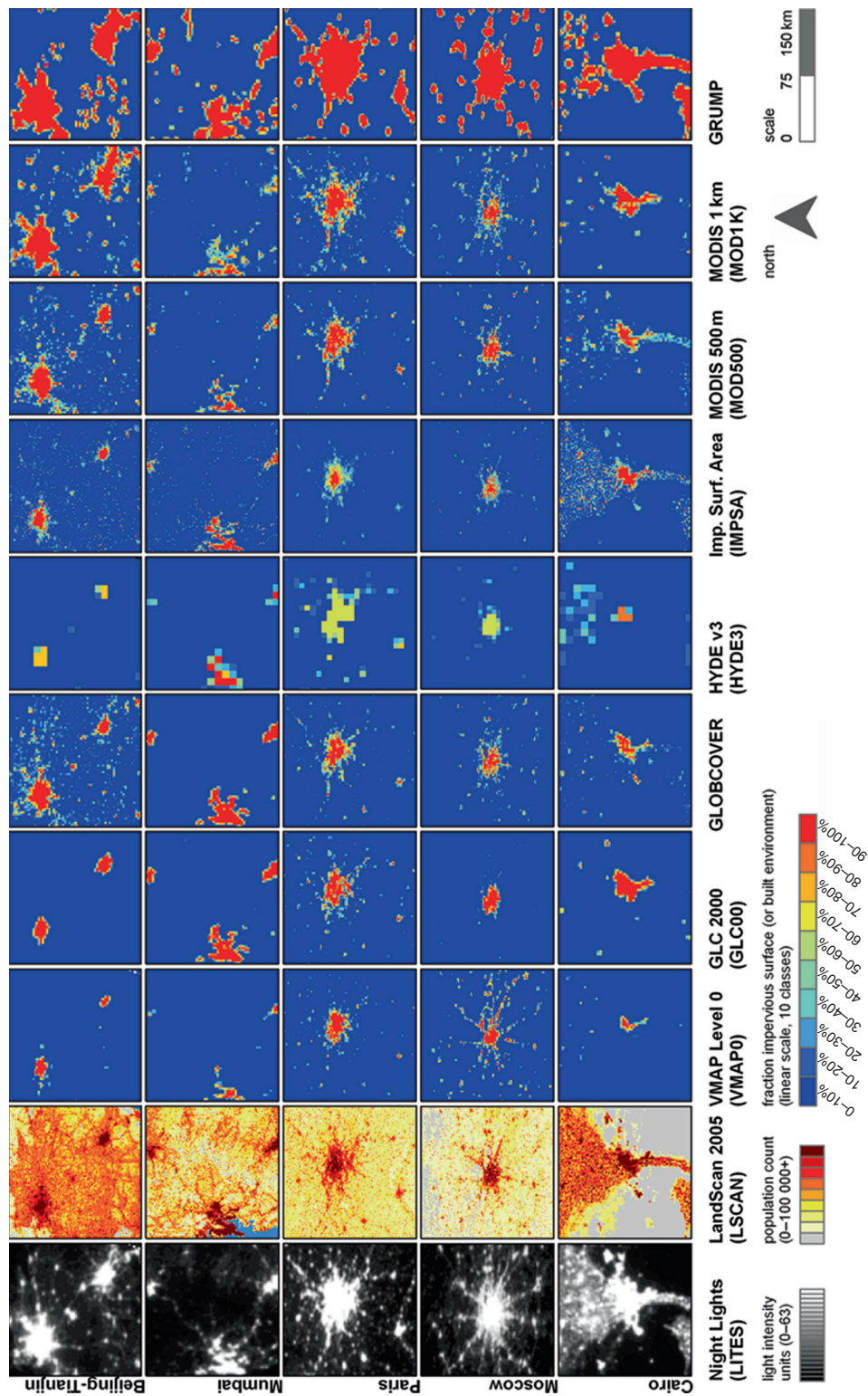


Figure 1.1: An inventory of first generation global urban footprint as reported by Potere et al. Figure from [19]

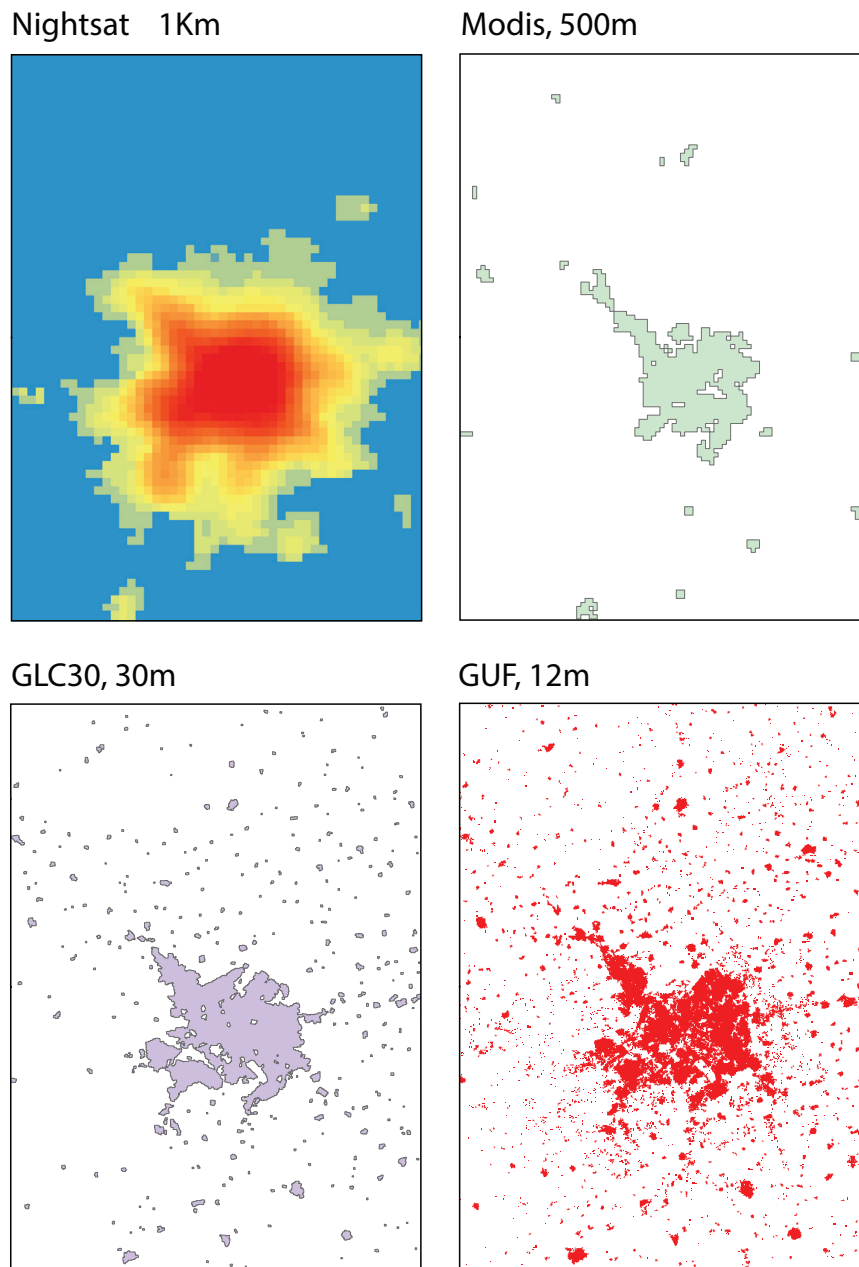


Figure 1.2: **First and second generation global urban inventories products.**A) The Nightsat at 1Km resolution, colours represent the average annual intensity of stable lighth. B) Modis 500m,C) the GLC30 at 30m and D) the GUF data at 12 meters resolution. It is possible to observe how low resolution data do not spot small settlements that might be an abundant part of urbanized area.

1.2 Urban scaling: Zipf's law, allometry, and the importance of being large

A pillar of quantitative urban geography is certainly the *hypothesis of urban scaling*, which, over the last fifty years [20], has lain at the very core of quantitative methods for understanding cities. In general terms, it postulates that all urban systems, at any scale of spatial aggregation, are similar. This hypothetically means that when looking at a map of an urbanized zone without a scale reference, which would indicate the size of the observed towns, one would be unable to determine whether one were observing a small village or a large town. The hypothesis of urban scaling also postulates the universality of this scale-free behaviour, which implies that urban patterns in China should be similar to those in Africa or Europe. This, in turn, implies that any process that governs urban evolution should be independent of local and initial conditions. Far from neglecting the uniqueness of each urban settlement, the urban scaling hypothesis highlights common patterns among 'dissimilar' cities. In a way, cities are considered as *variations on a theme* [45]. Urban scaling, similar to scaling in biology [46], can be distinguished into two main types: i) **intra-city scaling**, i.e., the distribution of an urban measurable, typically population or size, among cities and ii) **infra-city scaling**, i.e., a relationship between urban features within a city, for example, gross metropolitan product as a function of population.

Intra-city scaling will be addressed first, using the classical example of urban population ranking as explained by Gabaix [47] and Newman [48]. Suppose that, based on the populations of all of the cities in Europe, a graph is drawn in which the logarithm of the rank of each city is placed on the y axis (such as Rome I, London II, Paris III, etc.) and the logarithm of the population of each city is placed on the x axis. The series of points can be fairly well approximated by a straight line with a slope $\alpha = -1$, as shown in Fig.1.3. The implications of this apparently simple figure are very intriguing and have not yet been fully explained. It indicates that the probability of finding a city of a given size is a power law of the following form:

$$p(x) = Cx^{-\alpha} \tag{1.1}$$

The observation of a rank size rule that strictly follows $\alpha = -1$ has been addressed by George Zipf [49]; thus, this rule is also known as Zipf's law. Since Zipf's first publication, there has been a flurry of literature confirming the stability of Zipf's law across history [50, 51] and among various locations [52]. From a comparison of the literature, it seems clear that Zipf's law is a constant feature of all human settlements during any period and in any location. Reviews of this massive body of literature have been presented by Carrol [53] and, more recently, by Berry [54]. The simplicity, regularity and universality of Zipf's law have made it a fundamental prerequisite for any urban model. Many models have been proposed to explain Zipf's law for cities [55]. Thus far, the most reasonable, at least for urban patterns' analysis, have been proposed based on Gibrat's law [47]. Generally, these models operate in two phases: 1) Starting from a group of cities with random populations, they adopt a power-law distribution, introducing a scale-independent growth rate. This simply means that any force of agglomeration or growth acts on all settlements with the same intensity, not depending on the size of the town. Because this growth

is independent of scale, in the long run, it naturally generates a power-law distribution. 2) To force the exponent to be 1, the model forces the average size to be constant. Such multiplicative models are widely acknowledged as the best explanation of Zipf's law for cities, but they are also subject to several criticisms; for example, Batty recently noted their inability to simulate ranking volatility, i.e., they cannot predict or reproduce changes in the rank positions of cities [56]. Moreover, even if it is able to reasonably reproduce a power law with an exponent similar to the empirical one, Gibrat's law does not provide an explanation for the scale-free distribution of cities. In brief, it can be claimed that forces of agglomeration are operating at all scales, but why this is so cannot be explained.

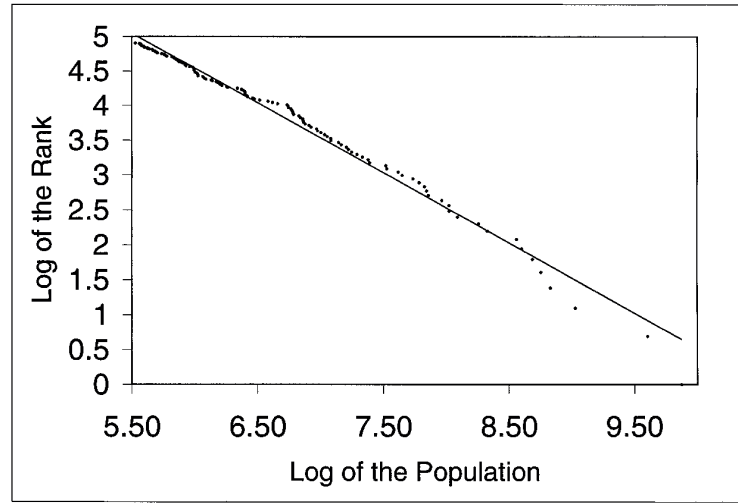


Figure 1.3: **Zipf's law.** Typical representation of Zipf's law as reported in [47]

However, cities are ultimately a spatial phenomenon, and if the same log-log plot is generated based on the sizes (rather than the populations) of objectively defined urban areas, a power-law distribution is again observed. Although the spatial correlation between an urban area and its population certainly cannot be linear (which would imply an anomalous constant urban density at all scales), in the literature, urban areas and populations are often confused. The scaling of urban areas pertains solely to the spatial extents and shapes of cities. In this sense, a crucial problem arises regarding the methodology applied to define city boundaries, which is, in a way, the same problem that affects global urban inventories, as described before. Nevertheless, regardless of the specific problems of urban boundary definition, which will be explored later, the main distinction between Zipf's law and urban land-use scaling lies in the mechanisms proposed to model them. Zipf's law has been mostly explained by means of multiplicative and space-independent models such as Gibrat's law, whereas the scaling of urban sites is spatially dependent and has been mostly modelled using gravity-like law models. Described below is a standard and effective gravity model proposed by Rybski et al. [57], which reasonably reproduces a power-law distribution of settlements. The approach here is to regulate the state of an urbanized area by representing the space in a discrete way. Consider a lattice space consisting of $N \times N$ empty cells W_i with coordinates $i \in \{1 \dots N, 1 \dots N\}$; after random seeding of this space with occupied locations,

the probability that an empty cell will be occupied is as follows:

$$q_i = C \frac{\sum_{k \neq i} W_k d_{i,k}^{-\gamma}}{\sum_{k \neq i} d_{i,k}^{-\gamma}} \quad (1.2)$$

where $d_{i,k}$ is the distance between cells i and k and $C = 1/\max(q_i)$ such that the maximum probability is 1. In the long run and for $\gamma > 0$, this simple model reasonably reproduces a power-law distribution. Because γ regulates the strength of attraction of an occupied site, a larger γ value results in a more centralized pattern, whereas a smaller γ value reproduces a more even and sprawled pattern, as seen in Fig. 1.4. Spatially explicit models of this kind offer a few advantages compared with non-spatial ones. They are able to simulate both the functional form of the size distribution and the geography of cities. They also explain urban growth as a sum of small and spatially discrete interactions, which appears to be a closer approximation of real urbanization dynamics.

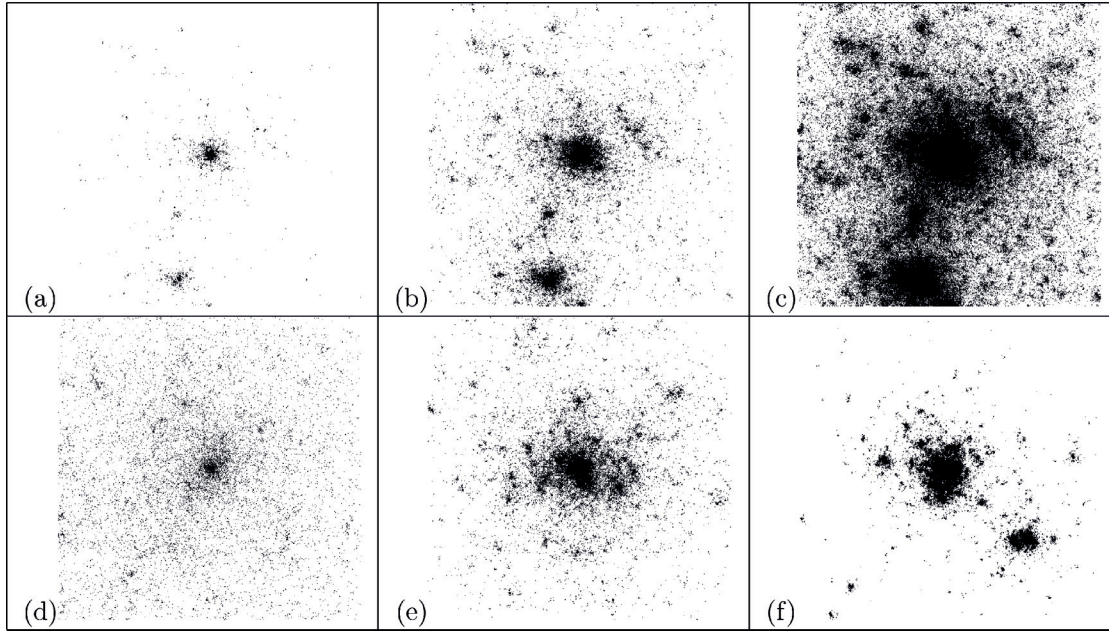


Figure 1.4: **Modelling Zipf's law.** Illustrative examples of model realizations for different iterations i and different exponents γ . (a)–(c) Different iterations of the model $i = 6, 10, 14$ ($\gamma = 2.5$, $N = 630$). Growth takes place close to occupied sites. (d)–(f) Realizations with different exponents $\gamma = 2.0, 2.5, 3.0$. The smaller γ , the more scattered the emerging structures are, the larger γ , the more compact they are. Image from [57]

Thus far, it has been observed that both the populations and sizes of cities scale with the same exponent, implying a relationship between these quantities. Linear relationships between urban features could be regarded as a general expectation; for example, per capita gross domestic product (GDP) should be expected to be independent of population. After all, the per capita GDP is the mean economic output of the inhabitants of a given city. Again, a population-independent

per capita GDP would imply a linear relationship between the total GDP of a city and its population. However, urban historians and urban economic scholars have noted unique behaviour in the economic performance of cities [30, 58, 31]. Paradigms such as *dynamic externalities* and *urbanization economies* are rooted in a combination of urban density and diversity, which promote both interactions (knowledge exchange) and economic competition, which, in turn, promote economic growth and innovation. Even if the related economic theories are focused on urban vs. rural economies, highlighting the primary role of cities in economic growth, they foster further understanding of urban economic output resulting from the interactions of individuals in a dense urban environment. Up to the present, important contributions in this direction have been made by Bettencourt, Lobo and colleagues [59, 45]. Such studies have revealed that the per capita GDP of a metropolitan basin, called the gross metropolitan product (GMP), depends on the size of its population through a power-law relation such as $GMP \propto Y population^\beta$, with β always being larger than 1 and close to 1.2 ± 0.02 . The GMP super-linearly increases with the population of a city, highlighting the fallacy of ranking cities by per capita quantities. Similar super-linear relationships have been also observed for patents, the number of new industries, and other urban-economy-related features. Thus, it has been proven that larger cities are wealthier than expected based on a linear assumption when compared with smaller cities. This observation offers a simple quantitative explanation of the increasing growth and expansion of cities, *i.e.*, a larger city offers an economically advantageous environment, and this is why large cities continue to attract inhabitants and keep growing. Yet, despite the great importance of these findings, the universality of scaling in cities is still under debate. Recent geographically sound studies have shown that the definition of urban boundaries is crucial for measuring the scaling of social and environmental performance with population size [60, 61]. Limited to the context of geography, these studies make a reasonable point. However, by focusing on the effect of the spatial definition of a city, they implicitly accept that scaling occurs only inside a metropolitan boundary, which is not so far removed from the core assumption of urban scaling. Thus far, the super-linear scaling of the economic performance of cities has been accepted as an empirical fact; however, it does not explain the consistently poor conditions of the majority of megalopolises in developing countries.

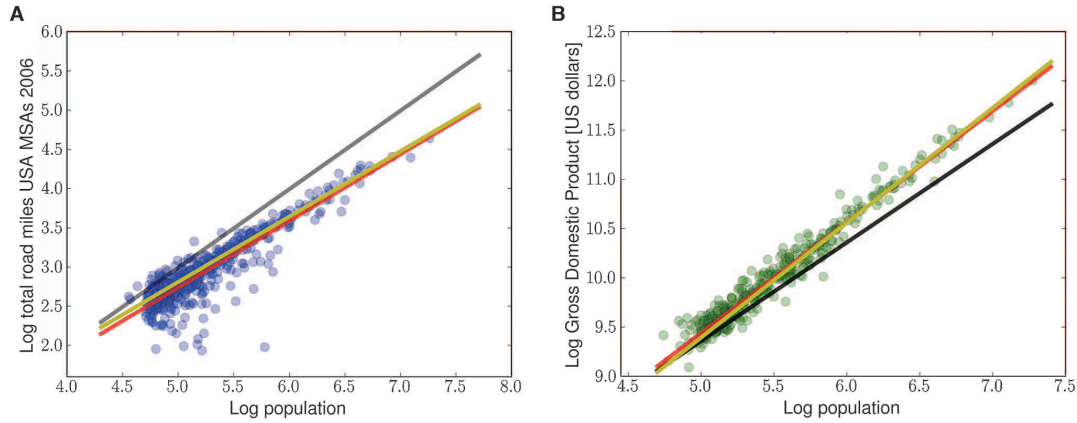


Figure 1.5: **Urban scaling.** Plot of sub and super linear scaling for US cities. In both plot the black line represents a linear correlation between X and Y variables. (A) Shows a sub-linear correlation between population and the total length of urban roads $population = Y_{roads}^\beta$ with $\beta = 0.84 \pm 0.03$. (B) A super linear scaling between population and the gross metropolitan product $population = Y_{GMP}^\beta$ with $\beta = 1.126 \pm 0.02$. Figure from [62].

The intrinsic scaling nature of cities has been extensively documented, but why is it still so important and puzzling? Several important questions that remain open despite the development of many models and theories are summarized below.

- **Range of city sizes.** City sizes, as shown in Fig. 1.3, span five orders of magnitude. In simple terms, this indicates that cities exist in a broad range of sizes. This, in turn, indicates that there is no preferred or optimal urban size, as might be expected. Why cities of all sizes exist remains an open question.
- **Scale-free urban landscapes.** A power-law distribution implies that the observed system is scale-free. Any power law of the form $f(x) = Cx^{-\alpha}$, if x is rescaled by a constant $f(x)$, remains proportional to $x^{-\alpha}$; thus, it lacks a natural scale and a characteristic unit. At any scale of observation, the functional form of the probability distribution is similar. This fact poses important questions concerning the very nature of large towns; is a large town simply a magnified copy of a small village, or is it totally different object? Such questions are still under debate. The meaning of self-similarity for cities, their dynamics and their mechanics have been the subject of many studies and are also a subject of this thesis.
- **Scaling range and lower cutoff.** The range of city sizes also pertains to the minimum and maximum of the range in which the power-law distribution is obeyed. Zipf's law for cities, as in many other biological and physical systems, deviates from the prevailing power law below some minimum value x_{min} , below which x goes rapidly to zero. This means that only large settlements follow the power law, whereas smaller ones follow a different functional form. This minimum bound can be attributed to two different possibilities: i) There may be problems with the resolution of the data such that below a certain value, they do not properly represent the observed phenomena. ii) If the data are considered to be correct throughout their entire range, then x_{min} might indicate a real boundary between two different systems;

for example, in the urban land use case, it might represent an objective method of defining different urbanization regimes, such as rural and urban development. However, in the case of cities, in terms of both urban land use and demographics, uncertainties in classification and resolution are typical problems. This precludes any clarification of the real nature of the minimum bound of the power law, i.e., the possible existence of a real boundary that defines urban and rural development remains to be clarified.

- **Maximum city size.** Although an average city size cannot be clearly defined in a power-law regime, a maximum can be clearly defined. However, the key factors that regulate the maximum city size are not yet clear.
- **Infinite mean city size.** A power-law distribution, depending on its exponent α , may have an infinite mean and variance. The mean $\langle x \rangle = C \int_{x_{min}}^{\infty} x^{-\alpha+1} dx = \frac{C}{2-\alpha} \left[x^{-\alpha+2} \right]_{x_{min}}^{\infty}$ becomes infinite for $\alpha \leq 2$. For a finite number of x values, as in the case of cities, this implies that the mean is not a well-defined measure; in brief, there is no meaningful average city size, suggesting that there is no optimal city size. It is important to remember that most empirical observations of the city size distribution indicate that $\alpha \sim 2$. It appears that in some sense, as noted by Krugman [55], the tension between urban economies and diseconomies (i.e., urban forces of agglomeration and spreading) result in the stabilization of the urban hierarchy across the singularity state of an infinite mean. However, the reason for such stability near 1 remains unclear and a topic of debate. Awareness of a possibly infinite mean and variance is of great importance for any landscape metrics, because this scenario means that in most cases, the observed mean and variance depend on the size of the observation and may not reflect any real urban geographic feature. However, this special characteristic of power laws has rarely been considered.
- **Hierarchical vs. optimal city size.** An important question regarding the city size distribution is whether it reflects, in any sense, the stability or maturity of a system of cities. Consequently, are early and non-mature urbanization patterns differently organized?
- **Analogy between nature and technology and self-organized cities.** Power laws appear in a broad range of phenomena in nature [63]. In fact, an important aspect of the scaling of urban systems lies in acknowledging, by analogy with other natural systems, the bottom-up emergent nature of urban characteristics versus the modern top-down planning approach to urban management [64]. This, in turn, poses several important questions regarding urban planning in the governing and managing of cities. If the final shape and organization of a city are inevitable and do not depend directly on any action, how we can use urban scaling observations to improve, manage and plan our urbanized world? This question, which pertains to the direct use of scaling theories and observations in the urban planning chain, is of great importance for the development of any scientific theory of cities. The deterministic and rational use of such theories also carries a few risks. A simple example is provided by a recent urban policy implemented by the Chinese government, which, based on a direct translation of the message that bigger is better (as claimed by many urban scientists) into urban development policy, is forcing migration from small and rural towns to large urban agglomerates. It seems that the debate surrounding the use of scaling approaches in urban management and urban planning is still in its infancy.

1.3 Complex spatial networks and urban form

In recent decades, a new approach to studying urban form, rooted in the physics of complex networks, has produced a series of new and important results. [65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75]. This approach is based on three main phases: i) modelling urban street networks as a planar networks, ii) producing a quantitative analysis of this network based on complex networks analysis and iii) interpreting the results as geographical or morphological attributes of cities. A simple example can clarify the application of this approach. Any old section of any European city is a product of a layering of urban plans, individual decisions, historical accidents, different styles of urbanism, transportation exigences and many other social factors. As a result, the final urban form is an intricate network of roads with many redundancies that, most of the time, is not easy to navigate. This is a common, everyday experience. By contrast, at the periphery of Detroit, for example, the pattern of roads is totally different. The top-down urban planning, the absence of any historical accidents and the prominently residential and single-house land use pattern have shaped a tree-like network structure, in which there are many dead-end roads, few network redundancies and very low connectivity between locations. It is evident that these urban street patterns are as different as Rome is different from Detroit. Urban street network analysis involves quantifying these differences and placing them in a scientific and reproducible framework of urban morphology in which typical qualitative information is corroborated with quantitative analysis.

A graph (or network) is a mathematical object which consists of two sets: \mathcal{N} and \mathcal{L} . The \mathcal{N} elements of the former are called *nodes*, while the \mathcal{L} elements of the latter (unordered couple of nodes) are called *links*. There are many ways to represent a graph, but the most common one is the *adjacency matrix* \mathcal{A} , a $N \times N$ square matrix whose entry a_{ij} ($i, j = 1, \dots, N$) is equal to one if link between nodes i and j exists and zero otherwise. Networks of street patterns belong to a particular class of spatial graphs called, *planar graph*, i.e. graphs whose links cross only on nodes. In the general and mostly used generalization method, the nodes represent street intersections, while the links are the streets centreline, a network made using this convention is called *primal road network*. A second method to model a network from a road pattern is based, with several different variations [65], on the geometry of roads and is called *dual road network*, in which nodes represent streets and links represent junctions. Figure 1.6 reports an example of the two generalization approaches. Road networks are weighted and each link (i, j) carries a numerical value w_{ij} expressing the intensity of the connection. The natural choice, in the case of roads, for the functional form expressing the weight of a link connecting nodes, say i and j , is to put w_{ij} equal to the length of the roads connecting, l_{ij} . Despite the increasing interest for urban roads network and the proliferation of studies, most of the results and findings can be categorized in two main classes. 1) measurements and distribution of network basic features, such as road's length distribution or node degree distribution, and 2) centralities of road networks. Below the most relevant findings have been reported, followed by considerations on potential gaps for further analyses.

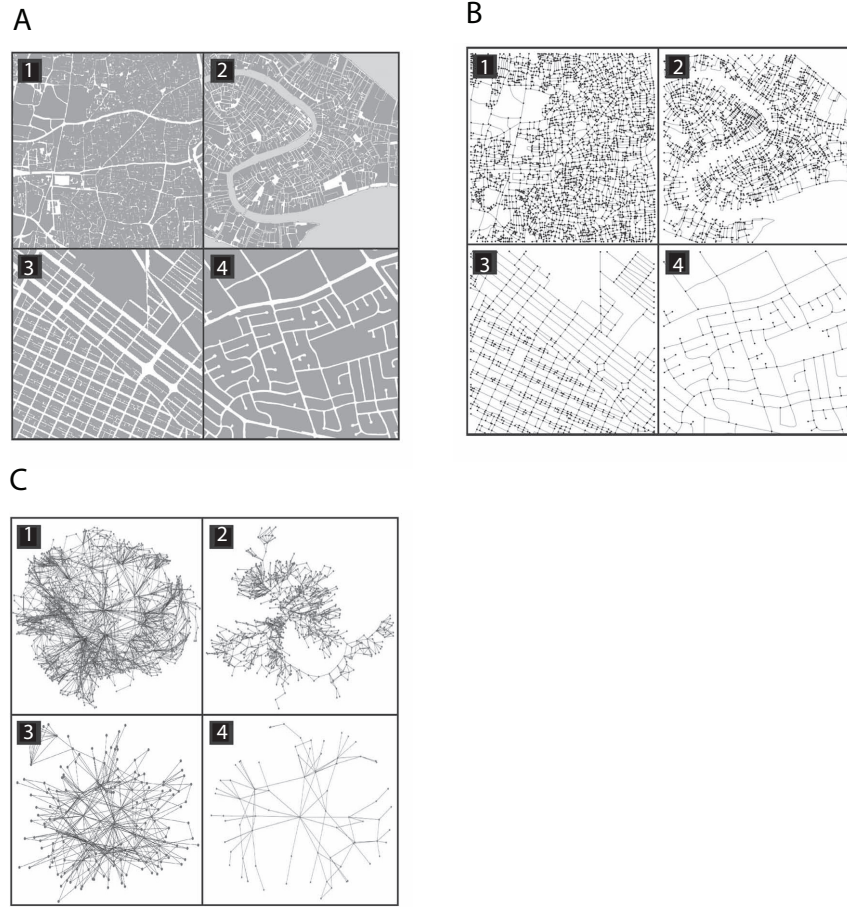


Figure 1.6: **From roads to networks.** A typical set of urban patterns in the top panels (A), and the corresponding primal networks in (B) and dual network in (C). Figure from [66]

Universal features of urban road networks vs. urban form.

There are some basic characteristics of urban road networks that are universal because they are constant among different regions and towns. The reason of such universality can be explained in term of network planarity constraints. Planar networks, by their physical nature, are all very similar to each other. However, variations within such constraints have been used to describe and classify urban roads patterns. For example, the *degree* of a node i , k_i , defined as the number of links incident with it, is always between 5 (in very rare cases) and 1 (in the case of dead end roads). Consequentially, the distribution of the *average degree* $\langle k \rangle = \frac{2E}{N}$, where E is the number of links and N the number of nodes, and that is nothing but the average of the degrees over all the nodes in the network, is very well peaked between 2.1 and 2.4. This regularity leads to an universal quantity $e = E/N = \langle k \rangle / 2$ that universally displays a value $1.05 < e < 1.69$ [76]. It basically shows that urban road networks display a form between a tree-like network ($e = 1$), and a 2d regular lattice ($e = 2$) as also reported in [70]. Fluctuations of the quantity e within this constraints tell us an important morphological feature about road networks. For example, in

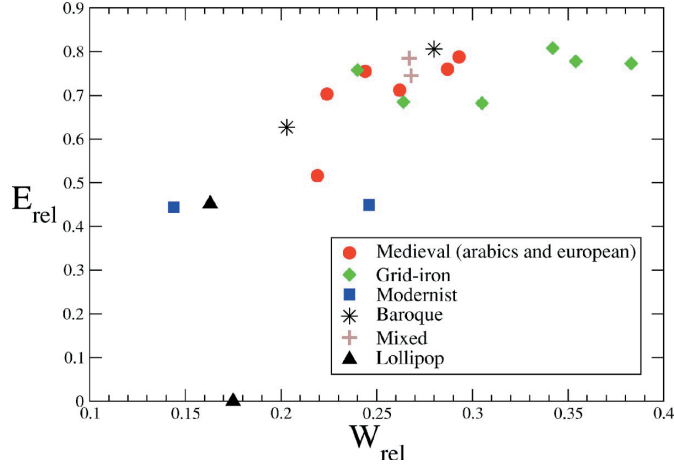


Figure 1.7: **Cost vs. efficiency describes historical road patterns.** The plot shows the Relative Efficiency E_{rel} as a function of relative cost W_{rel} as defined in [78]. It indicates a correlation between structural properties and a priori known classes of cities as medieval, grid-iron, modernist, baroque, mixed, and lollipop fabrics. Figure from [78] .

a recent study Masucci et al.[77] found different value of e can properly describe differences between UK and California roads. With the same purpose, an important and early study of urban road networks[78], proposed a classification of urban patterns given their similarity, in term of efficiency, to a tree or to a grid network. In this study efficiency has been measured as the average ratio between the euclidean and the path distance between all couple of nodes. Efficiency has been then computed over a set of historical (medieval, baroque, etc..) and contemporary road patterns. Results shows that, despite planarity constraints, historical accidents and urban planning are able to create different road patterns as shown in Fig.1.7. An other apparently stable characteristic of road network is given by the distribution $P(l)$ of roads' length. It has been speculated, in the case of London, that $P(l)$ follows a power law with a cut-off of the form $P(l) = l^{-\gamma}$ with $\gamma = 3.36$ [79]. In an other study [80], it has been shown that power laws distribution for roads length do not express any urban morphology features. In fact, roads' length distribution shows a more even distribution, similar to a log-normal distribution. The mean of road's length in the log-normal regime can be again used for a better description of urban form, as shown in Figure 1.8. It is clear that planarity cannot be used to claim about universality of urban roads network, It can rather used as a sort of null models to compare real road networks.

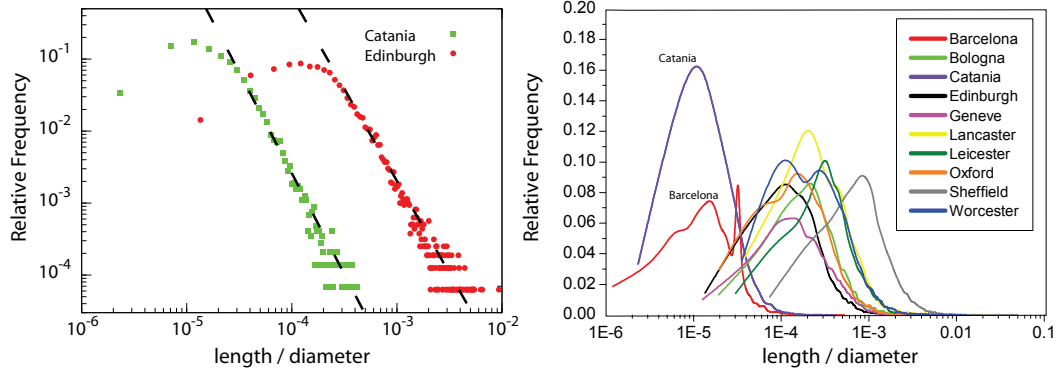


Figure 1.8: **Roads' length distribution of ten European cities.** Distribution of normalized length l for the ten cities in semi-log scale (right) and for two representative cities, Worcester and Catania, in the log-log scale (left). It is possible to observe how different visualization give different results, in the semi-log plot the distribution of the majority of streets describes different systems while in the log-log they have the same trend. Figure from [80].

Centralities in urban road networks

The concept of Centrality has been used since many years in network and social science and, starting from the seminal work by Wasserman [81], there has been a growth of literature both regarding centrality on social networks as well as other kinds of networks. Depending of the definition, centrality can be understood as meaning proximity between nodes, accessibility from other nodes, or being in a strategic position for connecting couple of nodes. From the seminal works by Lammer et al. [65] and Crucitti et al. [69] several studies applied centralities to urban road networks. It is clear that from different definitions of centrality, a node actor (a street junction) can be placed at different centrality ranks and that the same node can result with high value for a centrality while yielding weak values for another one. Below the definition and meaning of a set of centrality measures applied to street networks.

Betweenness centrality, C^B , is based on the idea that a node is more central when it is traversed by a larger number of the shortest paths connecting all couple of nodes in the network. More precisely, the betweenness of a node i is defined as in [81, 82, 83]:

$$C_i^B = \frac{1}{(N-1)(N-2)} \sum_{\substack{j,k \in \mathcal{N} \\ i \neq j, j \neq k}} \frac{n_{jk}(i)}{n_{jk}} \quad (1.3)$$

where n_{jk} is the number of shortest paths connecting j and k , while $n_{jk}(i)$ is the number of shortest paths connecting j and k and passing through i .

Straightness centrality, C^S , originates from the idea that the efficiency in the communication between two nodes i and j is equal to the inverse of the shortest path length, or geodesic, d_{ij} [84]. In the case of a spatial network embedded into a Euclidean space, the straightness centrality

of node i is defined as:

$$C_i^S = \frac{1}{(N-1)} \sum_{j \in \mathcal{N}, j \neq i} \frac{d_{ij}^{\text{Eucl.}}}{d_{ij}} \quad (1.4)$$

where d^{Eucl} is the Euclidean distance between nodes i and j along a straight line. This measurement captures to which extent the connecting route between nodes i and j , let's say between each street junctions, deviates from a virtual straight route.

The **Closeness centrality**, C^C , of a node i is based on the concept of minimum distance, in topological sense, i.e. the minimum number of edges traversed to get from i to j [85] and is defined as in [86, 87]:

$$C_i^C = \frac{1}{L_i} = \frac{N-1}{\sum_{j \in G} d_{ij}} \quad (1.5)$$

where L_i is the average distance from i to all the other nodes. Closeness centrality is a classical centrality index that has been widely used in urban geography and econometrics as well as in regional planning, where it gives an idea of the cost that spatial distance loads on many different kinds of relationships that take place between places, people, activities and markets.

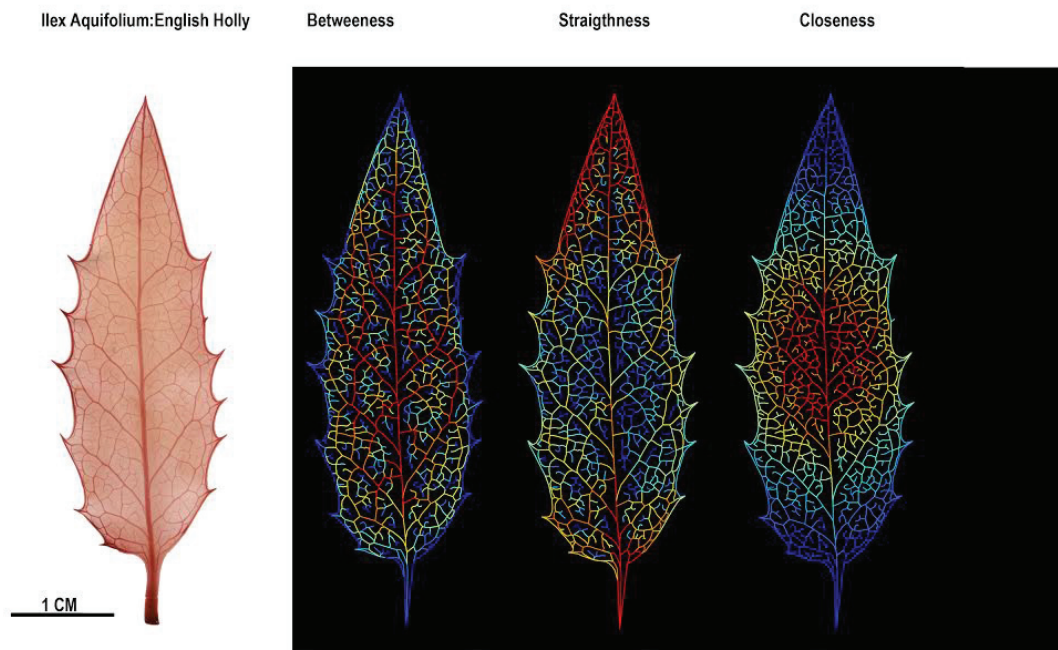


Figure 1.9: **Centralities in a leaf venation network.** Figure from the author.

Centralities have been used to classify urban road patterns: for example Crucitti and Porta [67] using the statistical distribution of Closeness, Betweenness and Straightness analysed a sample of twenty city parcels of one square mile. They found significant differences between cities and through a cluster analysis they proposed a classification of different urban patterns. The approach proved to be effective in capturing essential features of urban form as emerging in limited samples selected for their inner morphological consistency, but dealing with entire cities poses the problem of the classification of internally complex objects predominantly composed of different parts each possibly exhibiting different properties. Classification of roads network of entire city, and based on clustering algorithm applied to centralities, has been proposed by Strano et al. [80]. They show that cities belong to two main classes, those who have great geographical constraints, such as rivers or mountains, and those who growth without any geographical impediment. However, cities with great planning interventions, like for example Barcelona, belong to the first class, posing few questions about the role of big urban plans on the natural evolution of a town. Centralities have been also associated to social and economic urban features. An important example is given by the spatial correlation of centralities spatial displacement and micro-commercial activities. In [88], Porta et. al found that, roads with higher centralities, in particular Betweenness, are also roads with higher presence of micro-economic activities. Considering B_c as a proxy of traffic between all pair of roads, higher B_c implies more probability to be reached by users. In a way it as been provided evidence that the location of a shop is of great importance for it survival. Fig.1.10 shows results of this study.

Complex spatial networks applied to urban form analysis showed great potentials for a scientific understanding of cities. However, despite the number of emerging studies, this approach is still in its early stage and several problems need to be addressed. A clear example is given by the lack of analysis on large and very large road network that might encapsulate not only urban dynamics, but a wider set of land use conversion and dynamics. Other gaps are related to the limited number of case studies proposed in the literature. Recent increasing of computational power together with the availability of large and global road inventories, call for more analyses. An other evident gap is the absence of a multi-modal study. It is clear that urban networks are made not only by roads, and the effects of subways or rail networks on network measures can be of great impact.

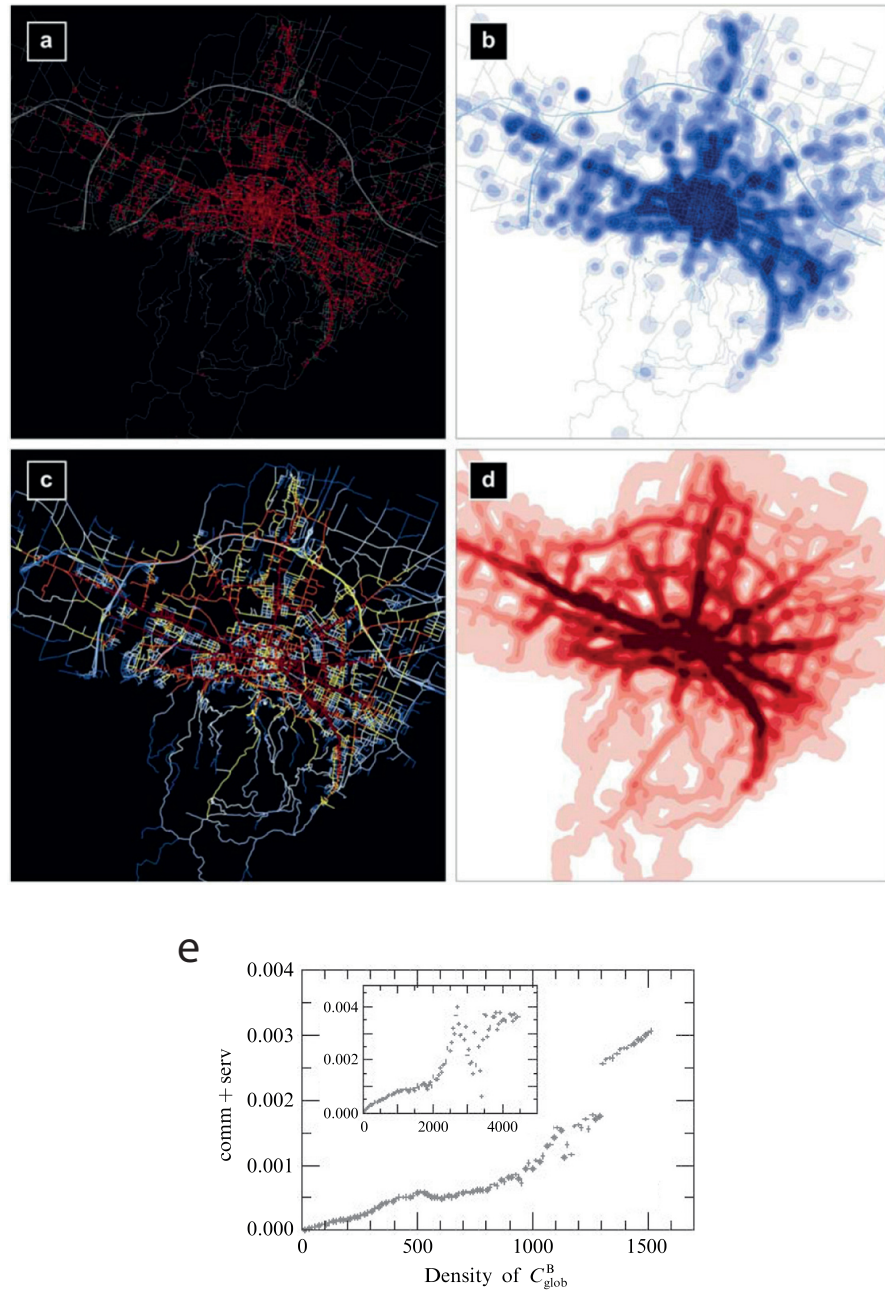


Figure 1.10: **Centralities density vs. commerce location in Bologna** a) location of commercial activity in Bologna. b) density of commercial activity defined with a Kernel density operator. d) map of Betweenness centrality for the road networks. e) density of Betweenness centrality. f) plot of commercial activity density vs. betweenness centrality density, each point of the plot represents a 10x 10 meters area as in b and d. Correlation in e shows that more Bc density more likely to find a commercial activity.

1.4 Main gaps and open problems

Different approaches to understanding cities bear their own privileges and constraints. The value of any scientific inquiry rests in the interdisciplinary efforts made to acknowledge it and, in a process of cross-pollination, recognize common problems and reciprocal solutions. This is the spirit of this thesis, which, with a unique focus on improving knowledge regarding urbanization, starts from a set of recognized gaps and solutions that originate from a multi-disciplinary view of urban studies. Below are reported and summarized the main perceived gaps and open questions that remain open and amenable to further exploration. Some of them are specific to a given field, but the majority come directly from the missing interactions between fields. All of them are explored, by means of various methodologies and case studies, in this study.

- **Lack of a global urbanization inventory** It is clear and well documented that at present, the state of human presence on Earth is not fully known. A very good approximation is available, but considering the extent to which human settlements drive global systems, better figures would be highly beneficial.
- **Scaling and global urbanization** The question of what are the implications of urban scaling with regard to the process of global urbanization remains to be satisfactorily addressed. Thus far, very few analyses have confronted this idea, and those that have done so have not progressed beyond statistical descriptions. Moreover, all of these analyses have used data of very low resolution, meaning that their results and speculations are of little practical use.
- **Defining urban boundaries and the unit of analysis** Despite the large debate surrounding it and the proliferation of different approaches, the problem of how to define city boundaries remains unsolved. The main question pertains to the existence of objective definitions of a town, a city, and a settlement that can be used to delineate the boundaries between different cities.
- **The universality of urban scaling** The universality of Zipf's law is widely accepted as universal and invariant. However, considering the very limited analysis performed thus far and the problem of the accuracy of the data (of any nature), there are no universal criteria confirming this universality. Therefore, this claim must be re-examined.
- **The universality of urban economic drivers** Is the virtuous loop between urban growth and economic growth universal? It has been universally accepted; however, the consistently poor conditions of rapidly urbanizing areas pose an important dilemma.
- **Transportation networks and urbanization** Despite the clear importance of transportation networks in the urbanization process, their mutual relationship thus far remains unclear.
- **Transportation network structures at large scales** All studies of urban network structures have been limited to urban-scale analysis. This means that any investigated properties of such networks do not reflect the entire structure of the road network, which naturally covers urban spaces but also spans the space between cities. An analysis of road networks at a larger scale is missing in the literature.

2 Material and methods

As stated in the previous sections, one of the most promising avenue to explore urbanization problems is routed in the interdisciplinary approach. This study has been developed at the intersection of three distinct scientific domains: 1) geographic information science including urban remote sensing, 2) complex spatial networks and 3) urban scaling. Both data and methods are coming from the three domains, below the entire data sets are reported. Each data sets have been used for a specific investigation as reported in Chapter 3. A detailed and more exhaustive description of each data, within its research context, is provided in Chapter 3.

2.1 Road and transportation networks

The study uses of several transportation networks from several sources. All of them needed to be corrected and prepared for the analysis stage. In general terms, the networks' preparation consists in a series of topological operations aimed to remove redundancies on vectorial data. Topological corrections can be executed under several environments. In this study a mix of Gdal libraries, Python scripts and Archmap functionalities have been used. After the topological corrections the adjacency matrices, which contain coordinates of source and destination and the length of network segments, have been extracted. In case of greatly uncorrected networks, like for example in the case of underground network from Open Street Map, there aren't any automatic operations or tool able to clean and correct the network. In this case, the only possible solution is to correct all the links and rebuilt the network manually. Below all the networks used for the study are presented.

- **Evolution of road networks in Groane location, Italy.**

To study the evolution of road networks several historical road networks have been extracted from historical cartography and areal photos. The area under study covers 125km² and includes 29 urban centers within 14 municipalities that have developed along two main radial paths, connecting Milan to Como and Milan to Varese. Seven street networks primal graphs at different times have been sampled, where the street junctions are represented as nodes and the roads (or streets) are the links of the networks. A mixed procedure of ArchMap tool extensions and python scripts operating over a geo-database has been used. Figure 2.1 reports all historical maps and materials. First, the street network has been drawn in vectorial format on the basis of a collection of historical areal images and historical maps

Chapter 2. Material and methods

Date	Source	Owner	Format
1833	Topographical Map of Lombardy-Venetia Kingdom	Italian Military Geographic Institute	Raster
1914	Map of Italy	Italian Military Geographic Institute	Raster
1933	Map of Italy	Italian Military Geographic Institute	Raster
1955	Aerial Photography Survey	Italian Military Geographic Institute	Raster
1980	Lombardy Regional Map	Lombardy Region	Raster
1994	Lombardy Regional Map	Lombardy Region	Raster
2007	Mosaic of Urban Municipalities Plans	Lombardy Region	Vectorial

Table 2.1: List of geographical information sources used to construct the Groane data set.

imported in the ArcGIS10 platform. Sources used for each temporal steps are reported in Table 2.1. The images have been imported as a .geotif extension, and all the layers has been projected using the coordinate system Monte Mario (Rome), Italy zone and Transverse Mercator projection. The first geo-referencing (alignment of different images) has been performed on the basis of historical buildings and landmarks such as paths and roads. The redraw operation was done using an ad-hoc python script for creating the node layers representing the street junctions. The second geo-referencing operation has been performed over the street junctions using the “spatial analyst tool” with the “spatial join” utilities provided in ArchMap10. Subsequently, for all time layers, using an ad-hoc python script we produced the weighted adjacency lists that has been used for the network analysis, the weight correspond to the real length of the street. After importing historical topographical and photogrammetrical data , the detailed road system has been constructed (including minor streets) at seven different points in time, $t = 1, 2, \dots, 7$, respectively corresponding to years: 1833, 1914, 1933, 1955, 1980, 1994, 2007. For each time, the associated primal graph has been built.

– Evolution of Paris network

By digitizing historical maps into a GIS environment, is reconstructed at five different time points: 1789, 1826, 1836, 1888, 1999. In particular, it is important to note that we have snapshots of the 3 street network before Haussmann works (1789-1836) and after (1888-2010). This allows to study the effect of such central planning. This data has been used in a precedent study [71]. The networks for 1789, 1826, 1836, 1888 are extracted from the following maps: 1789: Map of the city of Paris with its new enclosure. Geometrically based on the ‘meridienne de l’Observatoire’ and surveyed by Edffe Verniquet. Finished in 1791. 1826: Road map of Paris surveyed by Charles Picquet, geographer for the King and the duke of Orleans. 1836: Cadastre of Paris, Philibert Vasserot. Map constructed according to blocks and classified according to old districts. 24 Atlas, 1810-1836. 1888: Atlas of the 20 districts of Paris, surveyed by M. Alphand, and L. Fauve, under the administration of the prefect E. Poubelle, Paris, 1888.

– Other transportation networks

Oxford and Bologna road networks. The street networks of these two cities have been downloaded from Open Street Map [89]. They represent all streets including local urban streets and major roads. The boundaries of the networks were determined by the gradient of built-area density and include a minor peripheral zone of non urban streets.

Nantes water system. The Nantes water system was extracted from ‘eau-tuyau’, managed by Nantes Metropole. This dataset is copyrighted by Nantes Metropole and have been granted

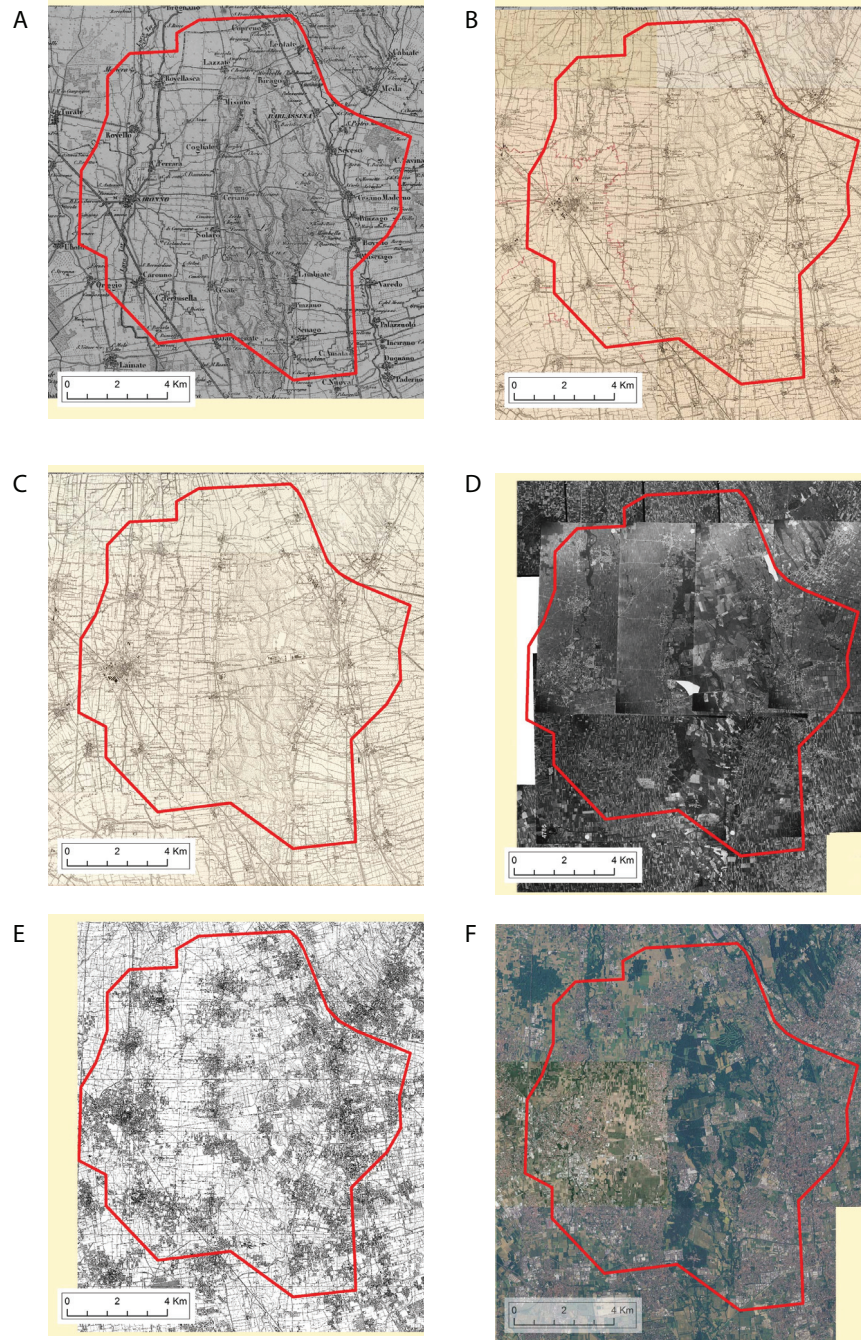


Figure 2.1: **Historical map sequences for Groane study.** A) Topographical Map of Lombardy-Venetia Kingdom, 1833 B) Map of Italy of the Italian Military Geographic Institute, 1914 C) Map of Italy of the Italian Military Geographic Institute, 1933 D) Aerial Photography Survey of the Italian Military Geographic Institute, 1955 E) Lombardy Regional Map, 1980 F) Lombardy Regional Map, 1994 G) Mosaic of Urban Municipalities Plans, 2007

to us for research purposes only.

Nantes roads network. It was extracted from the map 'BD Topo 2013' from the French national mapping agency. It is copyrighted by the IGN (Institut Geographique National).

Australian highways network. This network represents the full highway system of Australia and is part of the global road network from Data and Maps for Server provided by ESRI ArchGis under official ArchGis license. The data are ESRI copyrighted and have been granted for research purpose as described in the ESRI Proprietary Rights Acknowledgment-

UK rail network. This network represents the full rail transportation system in the UK. This dataset is in the public domain under the ODC Public Domain Dedication and Licence and has been downloaded from the website ShareGeoOpen [90].

- **Inventories of global road networks** The research explore the structural proprieties of the network composed by all major road on Earth. For this propose two products have been used: the gROADS [91] and the DeLorme World Base Map [92] data-sets. Both datasets contains major roads networks, however several important differences and discrepancies between them has been founded as is reported in Chapter.3.6. Below a description of data and the data source. 1: The gROADS global road network. The Global Roads Open Access Data Set, Version 1 (gROADSv1) was developed under the auspices of the CODATA Global Roads Data Development Task Group. The data set combines the best available roads data by country into a global roads coverage at 2009, using the UN Spatial Data Infrastructure Transport (UNSDI-T) version 2 as a common data model. It is an open and freely available at CIENISN website [91].

2: DeLorme World Base Map contains an updated version of all major roads on earth at 2014. The network is organized in four hierarchies: primary roads with limited access (H1), primary roads with non-limited access (H2), minor and secondary roads (H3), local roads (H4). The data do not contain urban roads. Fig. 2.2 shows a general overview of the DeLorme global road network. Topology of the road network has been corrected using Archmap software and had hoc Python scripts. The main purpose was to join connected roads at junctions with only two roads. A second purpose is to remove small road link representing highway ramps and cross roads intersections, thus not representative of any fragmentation process and potentially noise for statistical analysis.

- **Biological transportation networks**

In this research few biological transportation networks have been used. The main aim was to illustrate similarities and differences between man made and natural transportation networks. Biological and natural networks are original data, collected and extracted for this research. *Physarum Policephalum* This *Physarum Policephalum* plasmodium has been cultivated under sterile environment in a 10 cm radius Petri dishm Fig.2.3A. The Petri dish was kept in a dark at approximately 22 degrees for a period of 48 hours. Active plasmodium has been inoculated over a single food source composed by sterilized oat our tablets of 3mm radius. High resolution pictures were taken every then minutes. Five snapshots were chosen for analysis and their respective network were drawn on ArchMap software based on visual inspection. Only tubes thicker than 0.3 mm were considered. *Physarum polycefalum* is shown in Fig.2.3A.

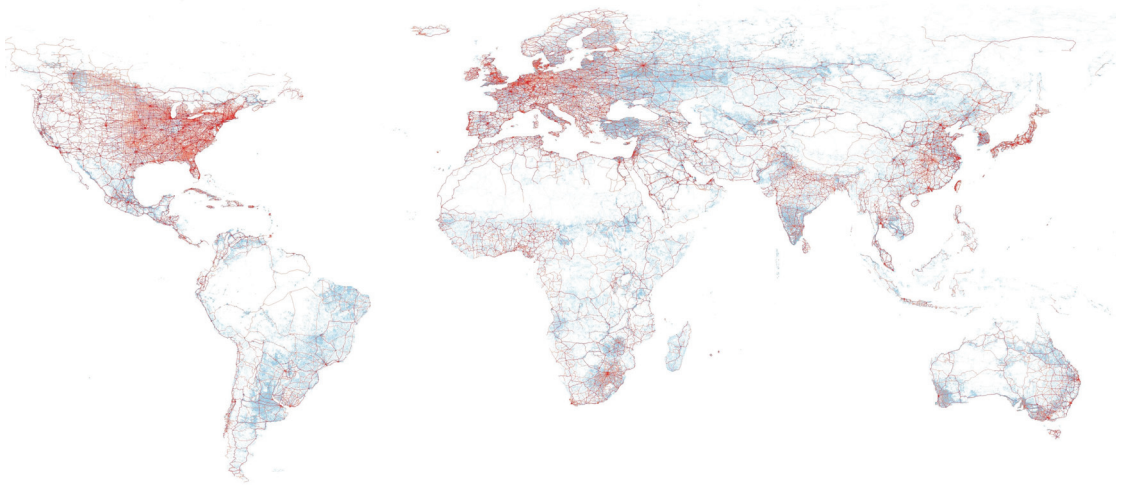


Figure 2.2: **A map of all roads on Earth.** Colours correspond to the hierarchy. Data from [92].

Leaves Ilex Aquifolium. It has been visually extracted using scanned imaged of cleaned leaf of *Ilex aquifolium*. The leaf is shown in Fig.2.3B.

Wing The vascular system of an anterior wing of a dragon was digitalized with a high resolution scanner. The insect was collected in the region of Aquasanta, Liguria (Italy), in June 2005. This dragonfly seems to belong to the family of Gomphidae. A picture is shown in Fig.2.3C

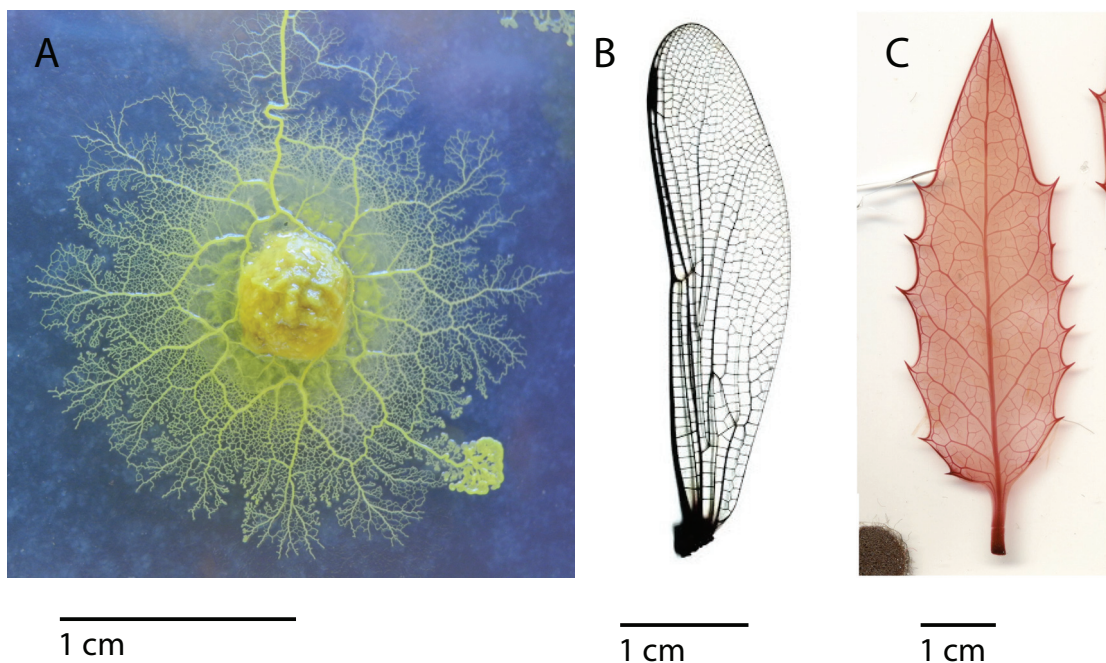


Figure 2.3: **Biological networks used in the study.** A) *Physarum poluchefalummm* B) dragonfly wings and c) *Ilex aquifolium* leaf.

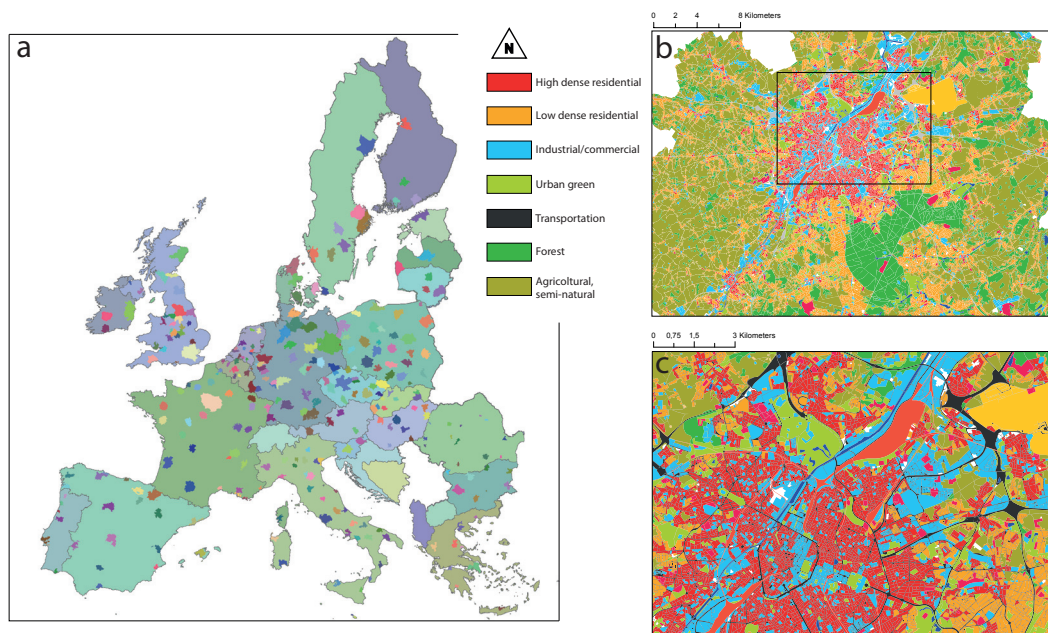


Figure 2.4: Location and map of the 300 European cities.

2.2 Land-use data

- **Land-use for 300 cities in Europe** Euro-atlas [93] consists in a pan-European comparable land-use vectorial data for urban areas and surroundings with more than 100.000 inhabitants in Europe. The spatial data on land-use have been produced and delivered by the European Environment Agency (EEA) [93] with the precise aim to provide an highly comparable and reliable set of land-use and land-cover data for large urban area to the wider scientific community. The EEA data-set provides 300 cities composed by vectorial geo-referenced sites divided in 19 land-use classes namely: continuous urban fabric (> 80%); discontinuous dense urban fabric (50% - 80%); discontinuous medium density urban fabric (30% -50%), discontinuous low density urban fabric (10% - 30%); discontinuous very low density urban fabric (< 10%); isolated structures; industrial, commercial, public; fast transit roads and associated land; other roads and associated land; railways and associated land; port areas; airports; green urban areas; sports and leisure facilities; agricultural; semi-natural and wetland areas; forest. The rationale behind the definition of the city boundary has not been properly clarified by the EEA nevertheless, exploring the data it is clear that they selected the urban zones based on principles of urban continuity and commuting. In principle, an urban zones have been defined by the continuity and proximity of artificial built up-area so that we include all of what can be called metropolitan areas. Fig.2.4 provides a map of locations and extent of the all the cities and a sample map of the city of Brussels.
- **The Global Urban Footprints, GUF** GUF is a binary settlement mask indicating built and un-built-up area at the very high spatial resolution of 0.4 arcsec (12m) and it covers potentially the global scale [43]. Classification method of built-up area makes use of the latest

radar satellite sensors called TerraSAR-X and TanDEM-X. Feature extraction is based on the analysis of local image heterogeneity [44]. Output of such classification represent the most accurate, but yet unpublished, global and uniform spatial repository of built-up area. It is important to emphasize here that the use of GUF at global level, since it is an unpublished and new data, represents an important element of innovation of the research.

- **Global Land Cover Dataset (GL30).** GL30 is a recent global land cover inventory provided by the National Geomatic Center of China [94]. It provides the global land cover classification for the year 2009 and at 30m resolution. The classification system includes 10 land cover types, namely cultivated land, forest, grassland, shrubland, wetland, water bodies, tundra, artificial surfaces, bareland, permanent snow and ice. Below a simple description of each class. In this study more than 300 scene for GL30 have used and composed in single and quasi-global mosaics. The analysed extent is shown in figure ref GL30A with an example of West Us mosaic. Despite the accuracy of land classification provided by GL30, in order to study urban areas, some manipulations of the data set were necessary. As shown in figure 2.6A, the land class artificial surface includes any kind of artificial surface, thus mixing roads surface together with built-up area. Image processing operation, namely shrinking and expanding, were applied to all the scene in order to remove long bodies such as roads and to separate urban patches, figure 2.6B shows the results of such operation.
- **Global crop land inventory** The 1 km global IIASA-IFPRI [95] crop-land percentage map for the baseline year 2005 has been developed by integrating a various of individual crop-land maps at global to regional to national scales. The individual map products include existing global land cover maps such as GlobCover 2005 and MODIS v.5, regional maps such as AFRICOVER and national maps from mapping agencies and other organizations. IIASA-IFPRI is a public data and it can downloaded from Geo-wiky platform at [96]. Figure 2.7 shows a global overview if such data.

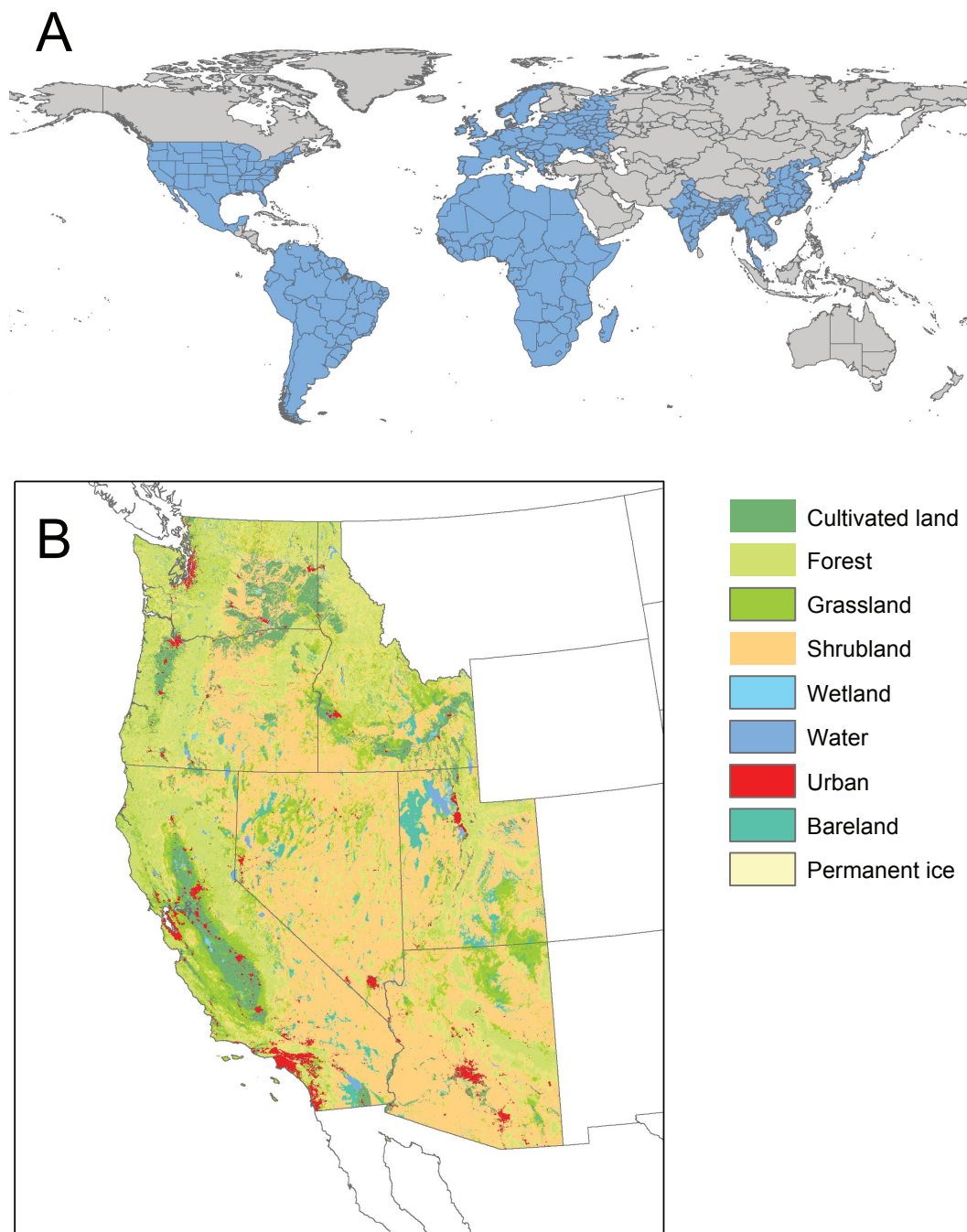


Figure 2.5: **Global Land Cover Dataset** A) In blue the extent of data from GL30 used in the research. B) A example of final mosaic composed for West US.

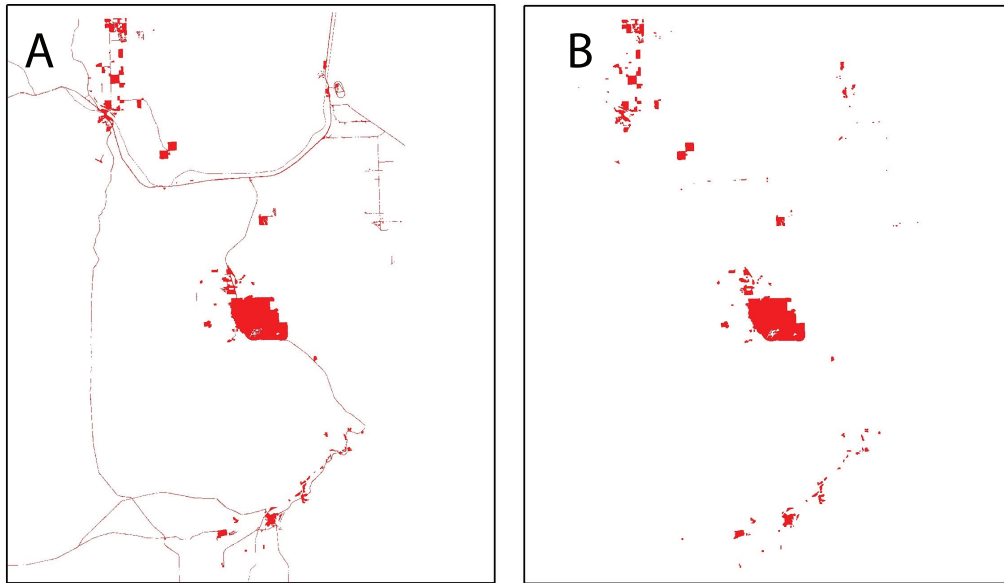


Figure 2.6: **GL30 preparation.** A) The original GL30 data for the land use class “artificial land”, it is possible to see the presence of roads. B) A corrected version of the previous image in which roads have been removed.

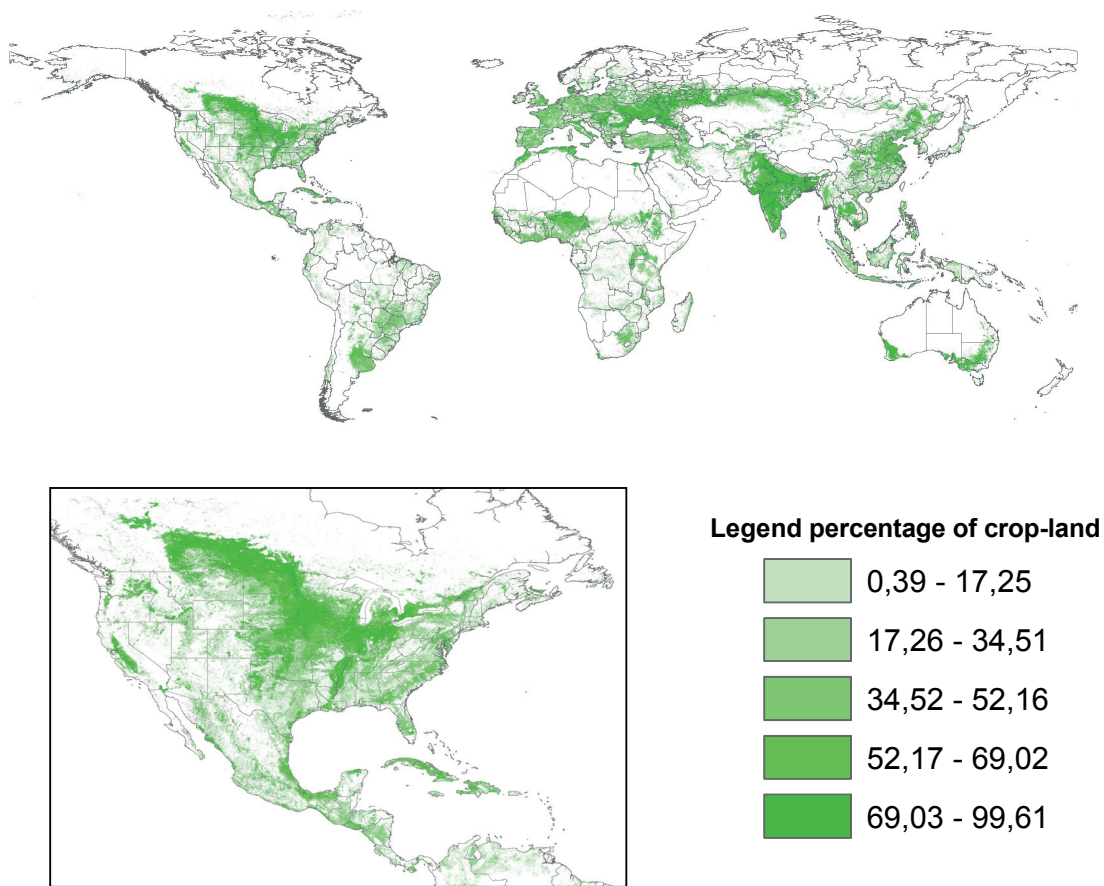


Figure 2.7: **Global crop-land spatial inventory.** The map shows the percentage of crop-land by Km^2 . Data from [95].

3 Results

Main gaps and open problems, as described in Sec. 1.4, concern the interaction and cross pollination of three main fields: remote sensing, complex spatial networks and urban scaling approach. Other gaps are related to the intrinsic difficulties in defining a proper scale of observation of urbanization and urban dynamics. Starting from these observations the proposed case studies try to cover both gaps, by improving interdisciplinary and covering the entire set of possible observation scales.

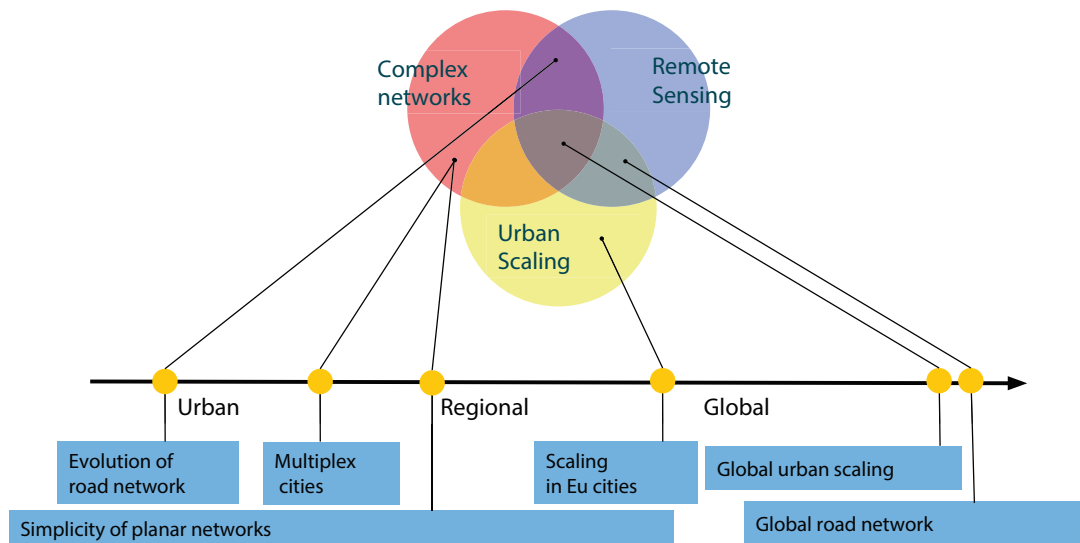


Figure 3.1: **Outline of the case studies** The case studies cover several scale of observation and are developed among remote sensing, urban scaling and complex spatial networks. The illustration shows at which interface each case has been developed as well as the scale of observation. The length of blue boxes indicates the scale covered by the case study.

Six case studies are proposed, each of them is a basis of scientific publications as reported at the beginning of each section. Case studies are presented in order of time and not following other specific rational. As illustrated in Fig. 3.1, the six case studies proposed in the thesis, moves from very local scale, or even microscopic scale, to global scale covering the entire globe. Fig. 3.1 also illustrated the

methodologies employed for each case. Each case is presented in form of scientific publication with introduction, methodology e data, results and conclusions.

3.1 Evolution of roads network, an empirical analysis

The study presents the analysis of an evolving road network in an area located north Milan. The time period covers more than 200 years. Results of this study have been proposed and published in the following scientific paper:

Strano E, Nicosia V, Porta S, Latora V, Barthélemy M, (2012) *Elementary processes governing the evolution of roads networks*. Nature Scientific Report 2.

Contributions of the candidate to the study include: research design, data production, data analyses, analysis of the results, writing the paper.

3.1.1 Introduction

As already illustrated, elementary spatial mechanisms that govern urbanization, leaving out specific historical, geographical, social and cultural factors, is nowadays more important than ever, especially because policy makers, professionals and researchers are actively looking for new paradigms in urban planning, land management and ecology [97, 98]. The existence of acultural and non-demographic drivers of urbanisation has been investigated using different approaches [20], including classical studies in regional sciences [99], theory of complex systems [100, 101, 102], urban theory [103] and remote sensing [104]. A consistent amount of literature in urban history and morphology indicates that roads are a fundamental driver in urban evolution and, at the same time, one of the long-lasting constituent elements of urban forms [105, 106]. However, a quantitative analysis of the historical development of urbanisation in metropolitan areas is still missing, and empirical evidence of the basic mechanisms governing urbanisation dynamics is still lacking. This section addresses this problem, and provides a study of the evolution of road networks over two centuries in a large urban area located north of Milan (Italy).

The area under study, known as *Groane*, covers a surface of 125 km², includes 29 urban centres within 14 municipalities, and has essentially evolved along two main radial paths, connecting Milan to Como and Milan to Varese, respectively. The former path was constructed by the Romans during the II century B.C., while the latter was created during the XVI century. In the last two centuries, the Groane area has faced a complex process of conurbation, changing from a polycentric region into a completely urbanised area. This conurbation process is common to many large European metropolitan regions. Despite local differences, it is possible to identify four distinct phases that characterize urbanization in the Groane area: i) *Rural phase*, (1800-1918): fundamentally pre-industrial, fully based on agricultural economy, with no major transportation infrastructures present. ii) *Early-urban phase* (1918-1945): between world wars period, still significantly based on agriculture economy, witnessing the first appearance of rail network, first small-scale sparse industrial colonisation, and limited expansion of rural settlements around the historical centres. iii) *Urban-industrial phase* (1945-1990): remarkable sprawled residential and industrial development (especially textile and mechanics)

along with population growth and highway construction. iv) *Metropolitan post-industrial phase* (1990-2012): decline of industrial activities, slower-paced urban sprawl, former polycentric organisation overwhelmed by the metropolitan continuum caused by the merging of expanded centres, increased long-range mobility due to the development of high speed trains and large highway systems [107]. It is important to note that this area has never been subjected to overall large-scale planning efforts, one reason being that 14 different administrative bodies preside over 14 different municipalities. In Fig. 3.2a is reported the road networks at different times, showing how the initial small separate villages have grown by the addition of new nodes and links, eventually merging together in an homogeneous pattern of streets. Let denote by $G_t \equiv G(V_t, E_t)$ the graph at time t , where V_t and E_t are respectively the set of nodes and the set of links at time t . The number of nodes at time t is then $N(t) = |V_t|$, while $E(t) = |E_t|$ is the number of links. By definition, we have $V_t = V_{t-1} \cup \Delta V_t$ and $E_t = E_{t-1} \cup \Delta E_t$, where ΔV_t and ΔE_t are, respectively, the set of new nodes and the set of new links added to the system in the time window $]t-1, t]$. In the following we will study the structure of the graph G_t at different times t , focusing in particular on the properties of the new links.

3.1.2 Results

Characterising the growth of the road network.

The Groane area is basically characterised by an uninterrupted growth in the period under consideration and displays remarkable modifications of the road system. As shown in Fig. 3.2c, in less than two centuries, the total number of nodes N has increased by a factor of twenty, from the original 255 nodes present at $t = 1$ (year 1833) to more than 5000 nodes at $t = 7$ (year 2007). However, the rate of growth is not constant over time: it is slow from 1833 to 1933, fast from 1933 to 1980, and slow again from 1980 to 2007. These different growth rates are the signature of distinct phases of the urbanisation process, and are strongly related to the growth of the population. The number of nodes N is a linear function of the number of people living in the Groane area (see Fig. 3.3a) or, in other words, that the average number of people per road intersection in the area remains constant over time.

These results highlight the peculiarity of an area where the population remains almost uniformly distributed in space, and no single urban center stands out over the others. In order to remove the demographic component and focus solely on the road network evolution, the number of nodes N has been adopted as a natural internal clock of the system in order to regulate the change of the network properties as a function of N . The number of links E grows almost linearly with N (Fig. 3.3b, top), showing that the average node degree $\langle k \rangle(t) = 2E(t)/N(t)$ is roughly constant over time, except for a slight increase from $\langle k \rangle \simeq 2.57$ to 2.8 when going from 1914 to 1980. In the bottom panel of Fig. 3.3b the total network length $L_{tot}(t) = \sum_{e \in E_t} \ell(e)$ (where e denotes a network edge and $\ell(e)$ its length) which increases over time as

$$L_{tot} \sim N^\gamma \tag{3.1}$$

where $\gamma \simeq 0.54$. This result is consistent with the evolution of almost regular two-dimensional lattices with a peaked link length distribution [108, 109]. Indeed, if we consider a uniform distribution of nodes with density ρ on a two-dimensional surface, the typical link length will be of order $\ell' \sim 1/\sqrt{\rho}$, which

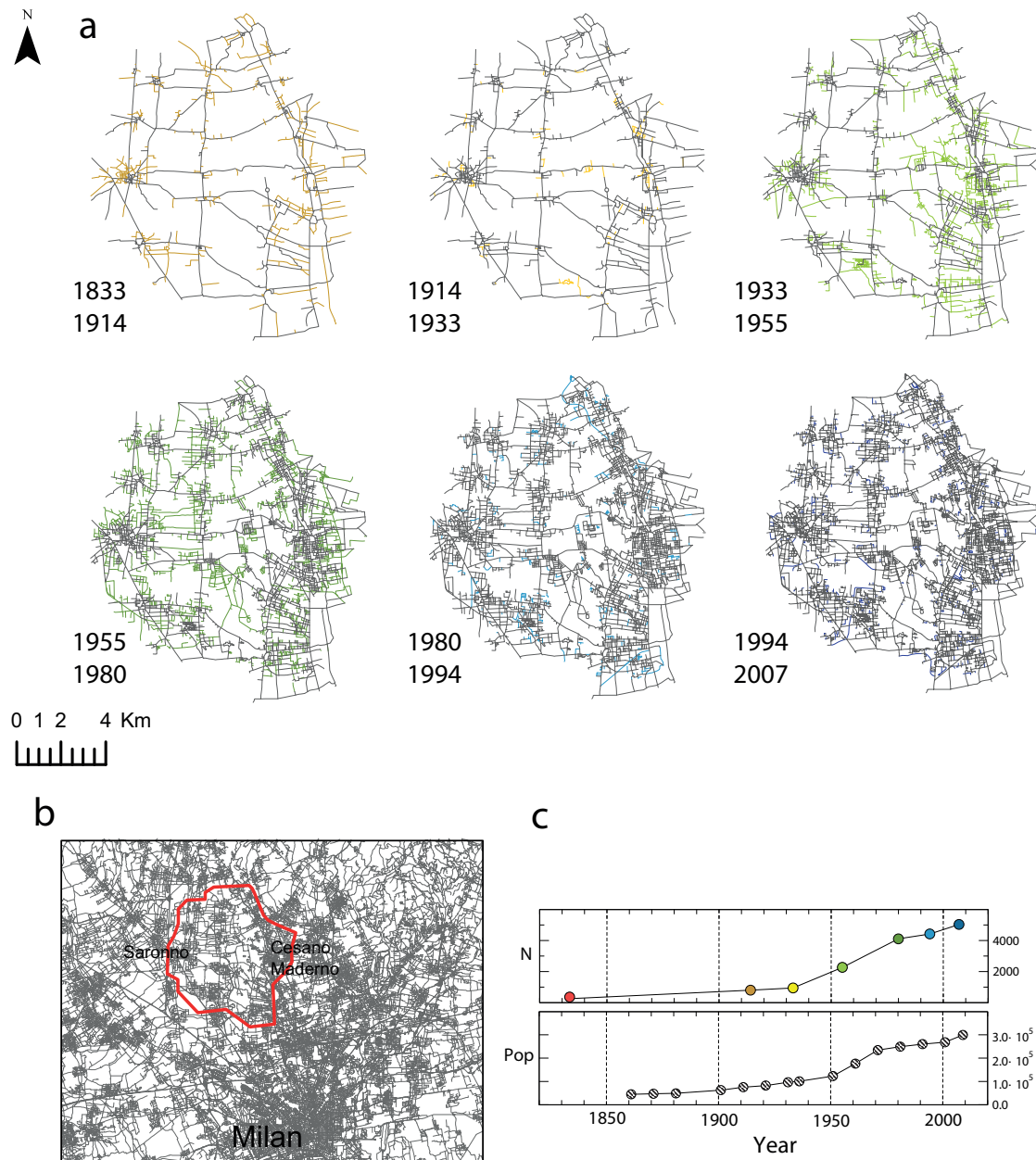


Figure 3.2: **Evolution of road networks** a) Evolution of the road network from 1833 to 2007. For each map in grey all the nodes and links already existing in the previous snapshot of the network, and in colors the new links added in the time window under consideration. b) Map showing the location of the Groane area in the metropolitan region of Milan. c) Time evolution of the total number of nodes N in the network and of the total population in the area (obtained from census data).

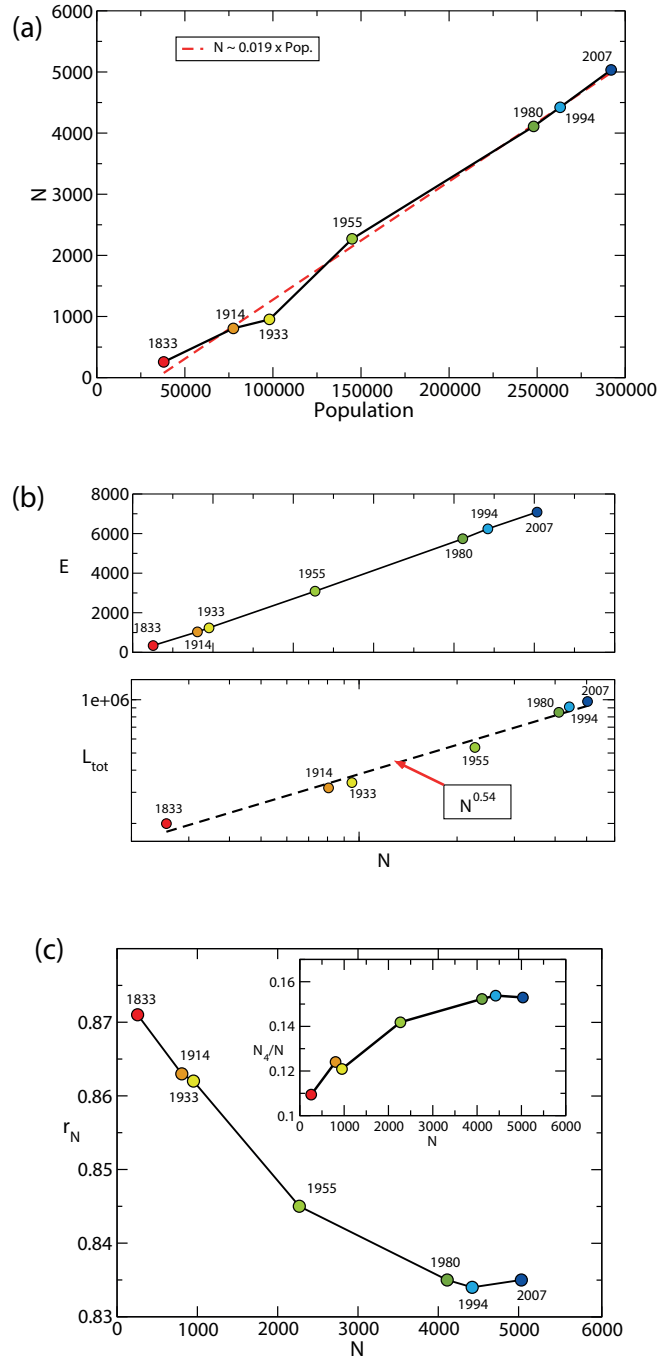


Figure 3.3: **Network evolution measures**(a) Number of nodes N versus total population (continuous line with circles) and its linear fit (red dashed line). (b) Total number of links E and total network length L_{tot} , as a function of the number of nodes N . The total network length increases as $N^{0.54}$. (c) Value of the ratio r_N between the number of nodes with degree $k=1$ and $k=3$, and the total number of nodes. In the inset the percentage of nodes having degree $k=4$ as a function of N . Notice that the relative abundance of four-ways crossings increases by 5% in two centuries.

implies that the total length grows as $L_{tot} \sim E\ell' \sim \sqrt{N}$, thus giving a value $\gamma = 1/2$.

Additional information on the structure of the road network can be obtained by looking at the quantity

$$r_N = \frac{N_1 + N_3}{\sum_{k \neq 2} N_k} \quad (3.2)$$

where N_k denotes the number of nodes of degree k (notice that we do not take into account nodes having $k = 2$ in the sum, since these are not usually considered as proper junctions). This quantity r_N measures the relative abundance of dead ends (corresponding to N_1) or T-shaped intersections (N_3), so that a small value of r_N indicates a dominance of $k = 4$ junctions and reveals the presence of a large amount of grid-like patterns. Conversely, the value of r_N is closer to 1 if the network has numerous T-shaped crossings and dead ends. The plot of r_N versus N (Fig. 3.3c) displays a steady decrease from $r_N \simeq 0.87$ at year 1833 to $r_N \simeq 0.835$ at year 2007, consistent with an increase of the percentage of four-ways crossings from $N_4/N \simeq 11\%$ at year 1833 to $N_4/N \simeq 15.5\%$ at year 2007. Recent studies have shown that the abundance of T-shaped crossings seems to be typical of self-organised or ‘organic’ urban networks, like those of Venice or Cairo [78], while the grid-like layout is particular to cities whose shapes are the result of large-scale top-down planning efforts [75], like Barcelona or New York. Consequently, one would conclude that what we are observing here is the evolution from an initial self-organised system to a rationally planned urban network. However, the Groane area at $t = 1833$ was actually a fully rural, not urban-organic network, and the successive evolution up to year 2007 has never witnessed any large-scale planning whatsoever. It is possible therefore to interpret the result of Fig. 3.3c as the signature of an evolution from a ‘pre-urban’ condition, with the dominance of dead-ends and 3-junctions typical of rural centres in the very early stages of growth (i.e. still constrained by the radial convergence of major roads), to an increasingly mature ‘urban’ state, in which the network expands on a relatively unconstrained land, further away from the dense radial system of original centres. At that later urban stage the dominance of the 4-junctions grid-like pattern does not result from large-scale planning, but from a piecemeal urbanisation, made of a multitude of small-scale, plot-based, scarcely coordinated developments in time, i.e. a substantially self-organised territorial order.

Evolution of cells: towards homogenisation.

Road networks are planar graphs consisting of a series of land cells surrounded by street segments. The statistics on the area and the shape of cells can be used to distinguish regular lattices from heterogeneous patterns. In particular, it has been recently observed that for Dresden (Germany) [65] and for a simple model of road networks [108], the cell area distribution $P(A)$ is a power law

$$P(A) \sim A^{-\tau} \quad (3.3)$$

with an exponent τ very close to the value 1.9. This value of τ can be explained in terms of a lattice constructed on a set of nodes with density fluctuations [109]. In Fig. 3.4a we report the distribution of the cell size in the Groane area at $t = 2007$, which is indeed a power law with the same exponent $\tau = 1.9 \pm 0.1$. However, the exponent changes in time, as reported in the inset: it takes a value $\tau \simeq 1.2$ at year 1833 and converges towards $\tau \simeq 1.9$ as the network grows. Because a larger exponent indicates a higher homogeneity of cell areas, we are thus witnessing here a process of homogenisation of the size

3.1. Evolution of roads network, an empirical analysis

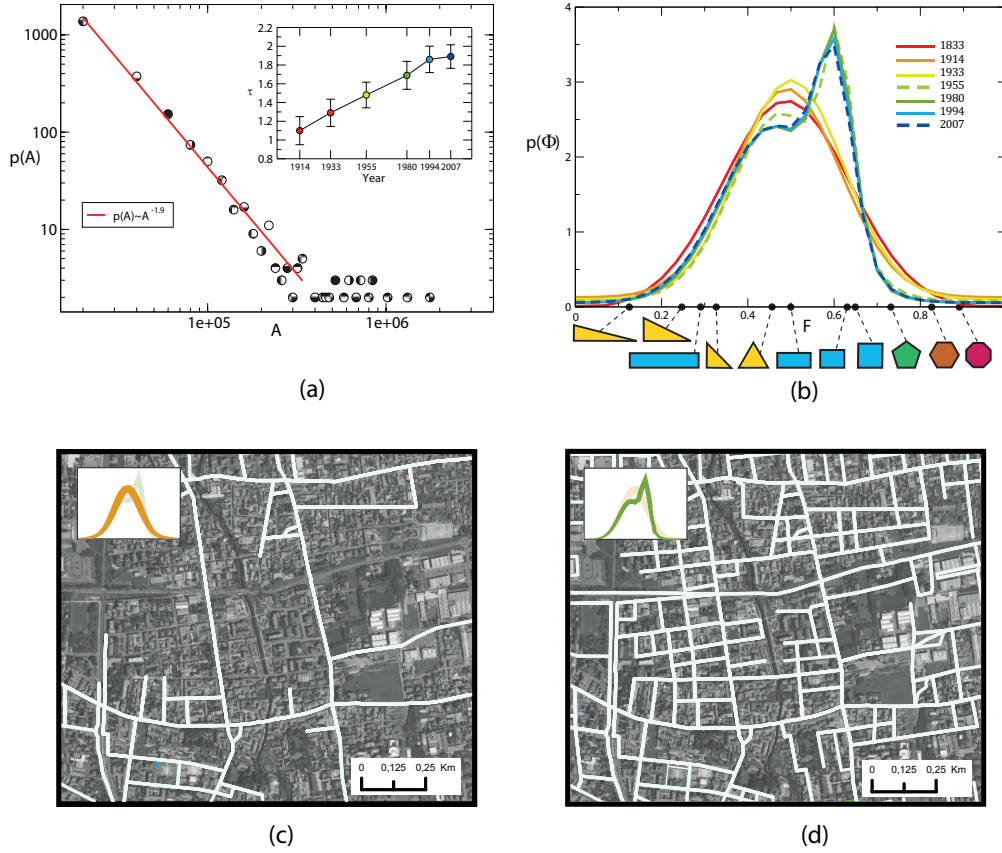


Figure 3.4: **Evolution of road cells** a) The size distribution of cell areas at $t = 2007$ can be fitted with a power-law $p(A) \sim A^{-\tau}$, with an exponent $\tau \approx 1.9$. The inset shows the value of τ at different times. (b) Distribution of cell shapes at different times, as quantified by the shape factor Φ . The shape factor of different polygons is reported at the bottom axis for comparison. (c) and (d) Maps showing the cell shapes (white lines) for the network as it is before 1955 (left panel) and as it is after 1955 (right panel). We see on the left panel that we have predominantly triangles and rectangles, while on the right panel we can observe a predominance of rectangles with sides of almost the same length.

of cells. Accordingly, the relative dispersion $\delta_A = \sigma(A)/\mu(A)$ of cell areas (where $\mu(A)$ and $\sigma(A)$ are the average and the standard deviation of A , respectively) decreases from 0.5 at year 1833 to 0.26 at year 2007, indicating that the variance of the distribution becomes smaller as N increases.

The diversity of the cell shapes can be quantitatively characterised by the so-called shape factor $\Phi = A/(\pi D^2/4)$, defined as the ratio between the area A of a cell and the area of the circle of diameter D circumscribed to the cell [110]. The value of the shape factor is in general higher for regular convex polygons, and tends to 1 when the number of sides in the polygon increases. The distributions $P(\Phi)$ reported in Fig. 3.4b clearly reveal the existence of two different regimes: before 1933 the distributions are single-peaked and well approximated by a single Gaussian function with an average of about 0.5 and a standard deviation of 0.25. Conversely, after 1955 the distributions of the shape factor display two peaks and can be fitted by the sum of two Gaussian functions. The first peak coincides roughly with the one observed before 1933, while the second peak, centred at 0.62, signals the appearance after 1955 of an important fraction of regular shapes, such as rectangles with sides of similar lengths. In

Fig. 3.4c-d we show the cell shapes at different times (before and after 1955) which visually confirm the findings of Fig. 3.4. The decrease in the relative dispersion δ_A and the increase in the fraction of regular shapes suggest that the network undergoes an evolution towards homogenisation. This appears to result from a combined process that exhibits two clear patterns: *i*) the fragmentation of larger cells of natural land into smaller ones, then heading towards urbanisation by medium-large manufactures or services, and *ii*) the mostly residential urbanisation of peri-urban natural land in successive rings around the historical urban centres. In particular, this latter pattern explains the emergence of more regularly shaped cells after the Second World War. As urbanization took place increasingly further from historical main roads, residential blocks became less constrained by the triangular shape defined by those roads and, as a result, increasingly followed a regular rectangular shape. This regular shape is in fact the most efficient way of subdividing land for urbanization, reflecting the inner organization of the block into equally regular rectangular plots. This grid-like layout tends to be applied extensively when local constraints, like main roads converging into a village, do not force the development along different patterns (for example T-shaped junction dominated).

Properties of new links: elementary processes of network growing.

Road networks grow by the addition of new streets (links) and new junctions (nodes). In Fig. 3.5 is shown the cumulative distribution of the length of new links according to the time-section in which they appeared first. The inset shows that the average length of new links steadily decreases over time, as expected from the general considerations of land fragmentation reported above. More precisely, we consider, at each time t , the length value $\ell_{90\%}(t)$ such that 90% of new links at time t are shorter than $\ell_{90\%}(t)$ (i.e. such that $P(\ell \leq \ell_{90\%}(t)) = 0.9$). We can notice that the value $\ell_{90\%}$ decreases in the period 1833-1933 from 625 meters down to 325 meters, while no sensible variation is observed from 1933 to 1994, even if the network keeps growing. In the last period, i.e. from 1994 to 2007, we observe another decrease of $\ell_{90\%}(t)$ from 325 to 225 meters. In addition, the relative dispersion of the length of the new links is almost constant and of order one, and the distribution does not vary too much after 1955. This phasing fits well with the historical development outlined in the introduction where during the rural and early-urban phase up to the second world war we observe the passage to an urban state that is then maintained along the core urbanisation age in the urban-industrial period until the 1980s, followed by a different post-industrial, metropolitan regime.

The nature of the growth process can be quantitatively characterised by looking at the centrality of streets. Among the various centrality indices available for spatial networks, we use here the betweenness centrality (BC) [111, 112, 66], which, as explained in Sec.3.2, is one of the measures of centrality commonly adopted to quantify the importance of a node or a link in a graph. Given the graph $G_t \equiv G(V_t, E_t)$ at time t , the BC of a link e is defined as

$$b(e) = \sum_{i \in V} \sum_{\substack{j \in V \\ j \neq i}} \frac{\sigma_{ij}(e)}{\sigma_{ij}} \quad (3.4)$$

where σ_{ij} is the number of shortest paths from node i to node j , while $\sigma_{ij}(e)$ is the number of such shortest paths which contain the link e . The quantity $b(e)$ essentially measures the number of times a link is used in the shortest paths connecting any pair of nodes in the network, and is thus a measure of

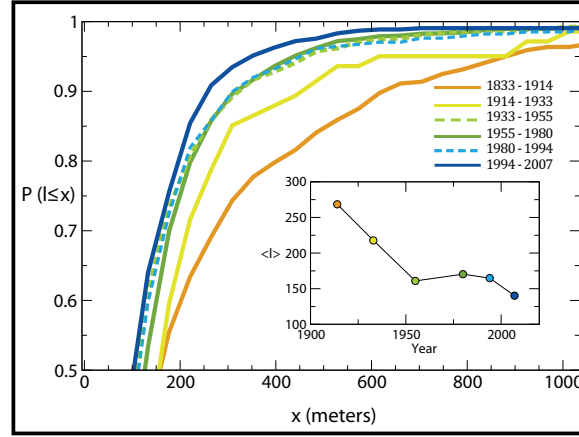


Figure 3.5: **Evolution of roads' length.** Cumulative distributions of the length of links added at different times. In the inset the average length. $\langle \ell \rangle$ of new links.

the contribution of a link in the organisation of flows in the network. In order to evaluate the impact of a new link on the overall distribution of the betweenness centrality in the graph at time t , we first compute the average betweenness centrality of all the links of G_t as:

$$\bar{b}(G_t) = \frac{1}{(N(t)-1)(N(t)-2)} \sum_{e \in E_t} b(e) \quad (3.5)$$

where $b(e)$ is the betweenness centrality of the edge e in the graph G_t . Then, for each link $e^* \in \Delta E_t$, i.e. for each newly added link in the time window $]t-1, t]$ we consider the new graph obtained by removing the link e^* from G_t and we denote this graph as $G_t \setminus \{e^*\}$. We compute again the average edge betweenness centrality, this time for the graph $G_t \setminus \{e^*\}$. Finally, the impact $\delta_b(e^*)$ of edge e^* on the betweenness centrality of the network at time t is defined as

$$\delta_b(e^*) = \frac{\left[\bar{b}(G_t) - \bar{b}(G_t \setminus \{e^*\}) \right]}{\bar{b}(G_t)} \quad (3.6)$$

The BC impact is thus the relative variation of the graph average betweenness due to the removal of the link e^* . We can measure this quantity for the different times-section and we report the results in Fig. 3.6. Remarkably, the distribution of $\delta_b(e)$ displays two well-separated peaks (Fig. 3.6d). The importance of the first peak tends to increase in time while the second peak decreases, until they mostly merge into one single peak in the last time-section (1994-2007). In order to understand the nature and the evolution of the two peaks, is possible to map the geographical location of new links with different BC impact.

In Fig. 3.6a-c are reported, respectively, the map of the network at year 1914, year 1994 and year 2007. Coloured in green the links whose centrality impact belongs to the left peak, and in red the links whose δ_b corresponds to the right peak. The new links are divided into two classes: green links (small δ_b , left peak) tend to bridge already existing streets, while red links (large δ_b , right peak) are usually dead-ends edges branching out of existing links, generating a new crossing and splitting the original link into two road segments. In order to better characterise the two classes of edges we compute $k_{min}(e)$, i.e. the

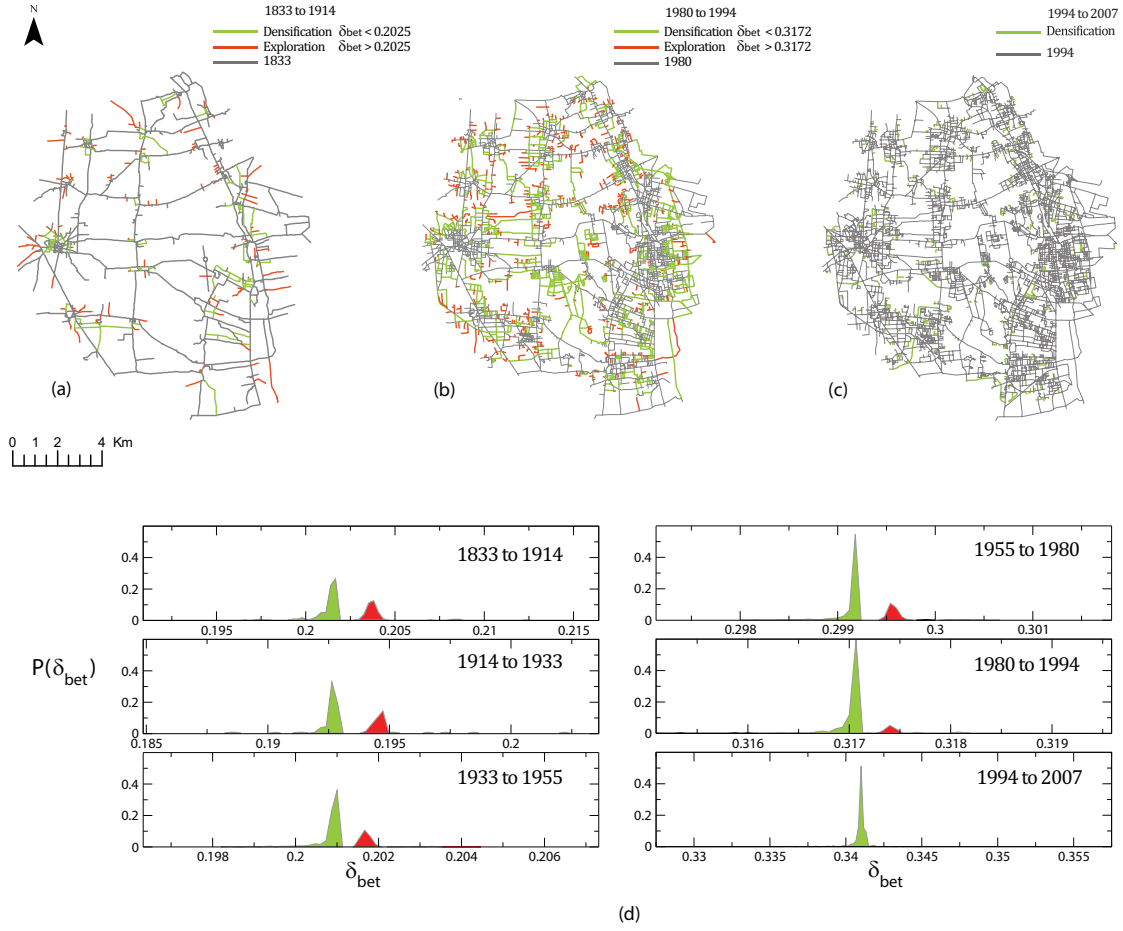


Figure 3.6: **Network growing phases.** The two phases of densification (green) and exploration (red), illustrated for the networks at year 1914 (a), 1994 (b) and 2007 (c). Panel (d) shows the probability distribution of the BC impact $\delta_b(e)$ for the different time snapshots. The red peak corresponds to exploration, and the green peak to densification. Notice that the red peak becomes smaller and smaller with time, and completely disappears in the last snapshot.

minimum of the degrees of both endpoints of a new link e , and we then consider, for each value of δ_b , the relative abundance of new links for which k_{min} is equal to 1, 2, 3, 4, respectively. The analysis of k_{min} confirms that the links belonging to different peaks have distinct qualitative and geographical features: links in the right peak have $k_{min} = 1$, while links in the left peak have $k_{min} \geq 2$. The distribution of BC impact thus suggests that the evolution of the road network is essentially characterised by two distinct, concurrent processes: one of *densification* (green links, left peak, i.e. lower impact on centrality, $k_{min} \geq 2$) which is responsible for the increase of the local density of the urban texture, and one of *exploration* (red links, right peak, higher impact on centrality, $k_{min} = 1$) which corresponds to the expansion of the network towards previously non-urbanised areas. At a closer sight, these two patterns tend in many cases to appear in a temporal sequence, the former acting in preparation for the latter, i.e. exploration being the first phase of a cycle of urbanisation then completed by a second phase of densification. Obviously, since the amount of available land decreases over time, at earlier time-sections (such as in the year 1914) the fraction of exploration is higher, while in the 80's it becomes smaller until it almost disappears in 2007. The only remaining peak mostly corresponds to densification with new links having $k_{min} \geq 2$.

Finally the relation between the age of a street and its centrality has been explored. In Fig. 3.7a are shown the links of the network in 2007 with a colour code depending on their age, and in Fig. 3.7b we report their BC. A simple visual inspection shows that highly central links usually are also the oldest ones. In particular, the links constructed before 1833 have a much higher centrality than those added at later time-sections. More precisely, the seven curves in panel Fig. 3.7c report the cumulative distribution of the BC computed on the network in 2007 for the links added at the different times. This is defined as the probability $P(b \leq x)$ that a link, appearing at a certain time-section, has a value of betweenness centrality b smaller than or equal to x in the final network in 2007. We can notice that the the historical structure of the oldest roads mostly coincide with the highly central links at year 2007. In particular, the inset of Fig. 3.7c indicates that more than 90% of the 100 most central links in 2007 (and almost 60% of the top 1000) were already present in 1833. This result reveals that a "backbone" of highly-central routes, that have framed the Groane area in the rural economy of the pre-industrial period, has been driving the development of the area across two centuries of industrialisation, urbanisation and de-industrialisation up to present days, without any major modification.

Urban morphologists have long observed that streets tend to persist in time much longer than other urban elements like buildings or land-uses [113]. This results confirm that at another level, that of the topology of the street network, the persistence of roads over time appears to be due to their multifaceted impact on the urban form at different scales by deeply informing factors like accessibility to space and resources, land ownership and land values, each of these involving different social actors and interests.

3.1.3 Discussion

This section reports a study of the evolution, over almost two centuries, of the street network in a large area close to Milan (Italy). Such an area is of interest for urban studies because it displays an important process of urbanisation. Urbanisation is reflected in the growth of the road network and displays different speeds at different times, with a fast increase occurring between 1933 and 1994. Results reveal a quantitative signature of urbanisation on the evolution of the shape and size of land

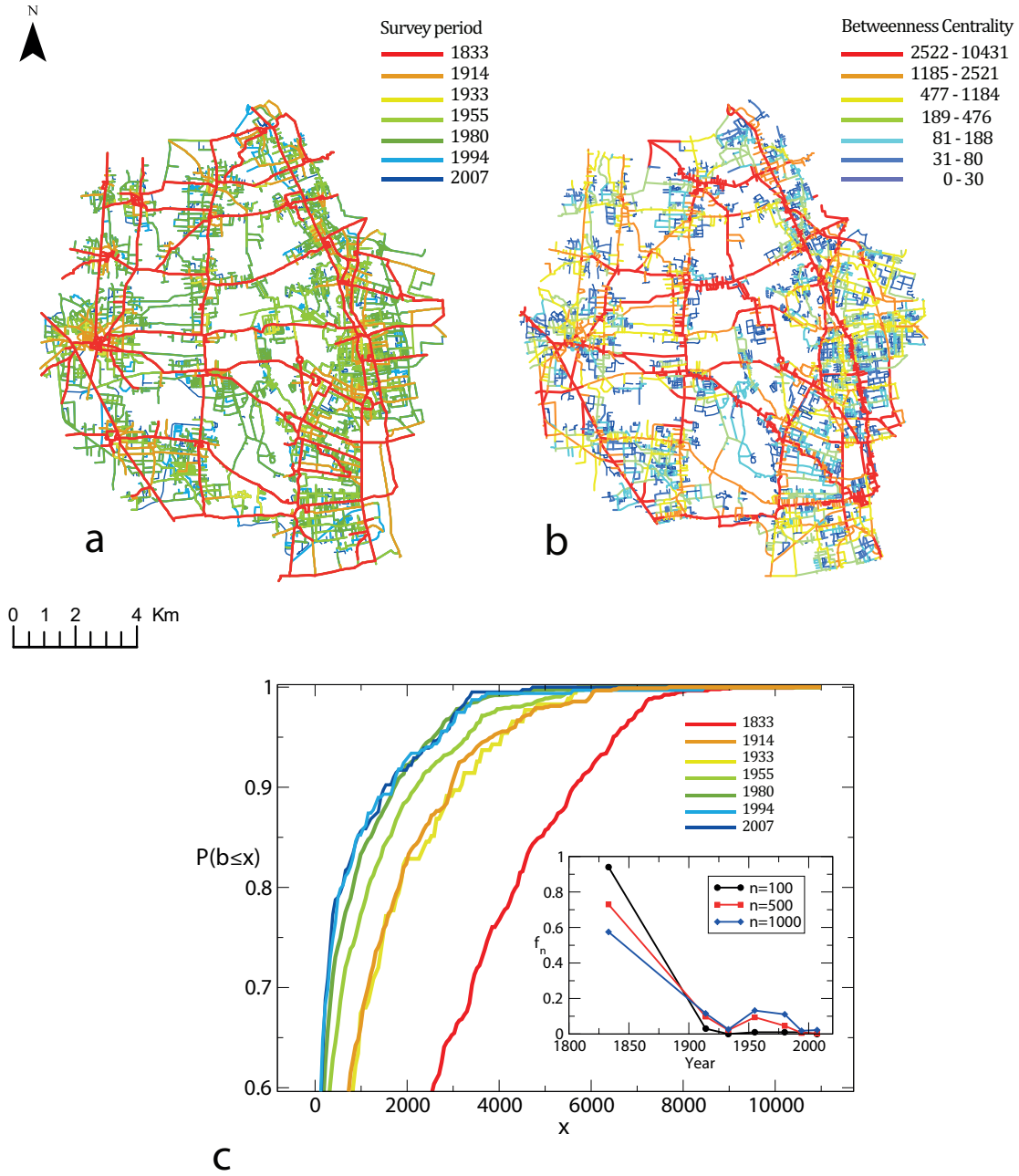


Figure 3.7: **Betweenness centrality vs. road's age.** Color maps indicating (a) the time of creation of each link and (b) its value of betweenness centrality (BC) at year 2007. (c) The cumulative distribution of BC of links added at different times. The inset reports the percentage of edges added at a certain time which are ranked in the top n positions according to the BC. Different curves correspond to $n = 100, 500, 1000$.

cells, which become more homogeneously distributed and square-shaped. Simultaneously, we observe a general trend towards a larger number of 4-way junctions, as opposed to an earlier structure of dead ends and 3-way junctions. These structural transformations appear to be the result of the interplay between two concurrent dynamics, namely densification and exploration. While exploration is typical of the earliest historical periods of urbanisation, densification predominates in the latest.

We were also able to quantitatively characterise the stability of the structure of most central streets over time. For instance, the most central streets in the network in 2007 largely coincide with the oldest ones. Central roads appear to therefore constitute a robust spatial backbone which remains stable over time, and characterises the evolution of the road system as a continuous expansion and reinforcement of pre-existing structures rather than as a sharp switch towards radically new configurations. The kind of evolution that we are witnessing in the Groane area is therefore certainly massive in overall quantitative terms, reflecting unparalleled changes in the pace and nature of economic, technological and social order during the transition from a mainly agricultural to a post-industrial age; however, it is still an organic form of evolution that builds at each step on the achievements of the previous, confirming in the long term a resilient structure that tends to persist in time all over the urbanisation process. These findings quantitatively support the hypothesis that spatial systems undergoing fine-grained forms of evolution tend to exhibit simple local dynamics of change, continuously expanding upon pre-existing structures, a morphogenetic behaviour that they appear to share with many living organisms and other natural systems.

The results of this research, including the existence of the two main dynamics of exploration and densification, together with the persistence over time of a structural backbone made of highly central streets, cannot be extended to the generality of urbanisation processes unless supported by further investigation of cases in different geographical and economic positions.

3.2 The simplicity of planar networks.

This study presents a new metric to analyse planar graphs. Such measure has been used for a comparative study of natural and man-made transportation networks. Results of the study have been published in the following scientific paper:

Viana M, Strano E, Bordin P, Barthelemy M, (2013) *The simplicity of planar graphs*. Nature Scientific Report 4.

Contributions of the candidate to this study include: design of the research, data preparation, data analysis, analysis of the results and writing of the manuscript.

3.2.1 Introduction

As described in Sec.2.1 urban and transportation networks are quasi-planar networks. A planar network is a graph that can be drawn on the two-dimensional plane such that no edges cross each other [114]. Beyond urban models, planar graphs pervade many aspects of science and they are the subject of numerous studies in graph theory [115]. Planar graphs are also central in biology where they can be used to describe venation patterns of leaves or insect wings. In particular, the vascular network of leaves [116, 117] displays an interesting architecture with many loops at different scales, while in insects, the vascular network brings strength and flexibility to their wings. In city science, planar networks are extensively used to represent, to a good approximation, various infrastructure networks [109]. In particular, transportation networks [110] and more recently streets patterns, as described in Chapter.1.3.

However, despite a large number of studies on planar networks, there is still a lack of global, high-level metrics allowing to characterize their structure and geometrical patterns. Such a characterization is difficult to achieve and in this section an important aspect of planar graphs is investigated, which is intimately connected to their geometrical organization. In this respect, new metrics are defined and tested on various datasets, both artificial (roads, highways, railways, and supply networks) and natural (venation patterns of leaves and wings, slime mould) enabling us to obtain new information about the structure of these networks.

First static networks (see Fig. 3.11) have been analysed. The streets of cities (Bologna, Italy; Oxford, UK; Nantes, France), the national highway network of Australia, the national UK railway system, and the water supply network of central Nantes (France). In the case of biological networks, we study the venation patterns of leaves (*Ilex aquifolium* and *Hymenanthera chatamica*), and of a dragonfly wing. Details on these datasets can be found in the Chapter 2.1.

Three datasets describing the time evolution of networks at different scales have also been considered (see Fig. 3.12): at a small scale and in the biological realm the evolution of a slime mould network is analysed. At the city scale the road network of Paris (France) from 1789 until now. Paris was largely transformed by a central authority (the prefect Haussmann under Napoleon III) in the middle of the 19th century and the dataset studied here displays the network before and after these important transformations, offering the possibility to study quantitatively the effect of top-down planning [71]. At

the multi-town level, we study the road network of the Groane area (Italy). These networks allow to explore different systems at very different scales from 10^{-3} (Slime mould) to 10^6 (Australian highways) meters.

3.2.2 Material and methods

A new metric, the simplest path

Generally speaking, is possible to define different types of paths for a given pair of nodes (i, j) . A usual quantity is the shortest path of length $\ell(i, j)$ which minimizes the distance travelled to go from i to j . It is possible however another path which minimizes the number of turns - the simplest path, of length $\ell^*(i, j)$ (if there are more than one such path we choose the shortest one). Fig. 3.8a displays an example of the shortest and simplest path for a given pair of nodes on the Oxford (UK) streets network. To identify the simplest path, we first convert the graph from the primal to the dual representation, where each node corresponds to a straight line in the primal graph. These straight lines are determined by a continuity negotiation-like algorithm. Edges in dual space, in turn, represent the intersection of straight lines in the primal graph (see Fig. 3.8b).

All the simplest paths of a given network were calculated in the dual space by converting the networks from the primal to the dual representation, where straight lines are mapped into nodes and the intersection between straight lines were mapped into edges. Straight lines are found by using a version of the *ICN* (Intersection Continuity Negotiation) algorithm [66]. More specifically, given an edge (i, j) , we search among the adjacent edges attached to j , (j, k) , that one that is most aligned to (i, j) . If the angle $\theta_{i,j,k}$ between (i, j) and (j, k) is smaller or equal to $\theta_c = 30^\circ$, we assume that these two edges belong to the same straight line. The value of $\theta_c = 30^\circ$ is an arbitrary value, however if applied to all the networks it not represents a systemic biases. This procedure continues until no more edges are assigned to the same straight line. Then, the procedure is repeated in opposite direction starting from the adjacent edges attached to node i . Once assigned to a straight line, an edge is removed from the network. As it is, this algorithm produces different networks depending on the choice for the initial edge. To overcome this ambiguity, the algorithm always starts with the edge that give us the longest straight line for a given network. After this straight line is fully detected and its edges deleted, we choose the next edge that will give the second longest straight line and so on. The algorithm ends when there are no more edges left in the network.

Once all straight lines have been identified, the dual representation is built by looking at the intersection between straight lines. Each straight line is mapped onto a node in the dual space and two nodes are connected together if their respective straight lines intersect each other at least once. To illustrate this process, we show in Fig 3.8(b,1) an example of planar network in the primal representation where the edges are colored according the *id* of the straight line they belong to. In Fig. 3.8(b,4) is shown the dual representation of the same network. It is important to note that the longest straight lines, in this example represented by orange, red, green and magenta give rise to hubs in the dual space.

In order to calculate the simplest path between nodes 1 and 2 from Fig. 3.8(b,1), the algorithm searches for the shortest path between their respective straight lines in the dual space, cyan and blue in this case. As it can be seen, there are two paths with length 4, A and B . Each of them define a subgraph in

the primal representation, here represented by the set of magenta lines in Fig. 3.8(b,2) and red lines in Fig. 3.8(b,3) for paths A and B, respectively. Then the shortest path distance between nodes 1 and 2 over these subgraphs has been evaluated and we adopted the shortest one as the simplest path - green dashed in Fig. 3.8(b,3). The black path in Figs. 3.8(c,d) represents the shortest path between nodes 1 and 2.

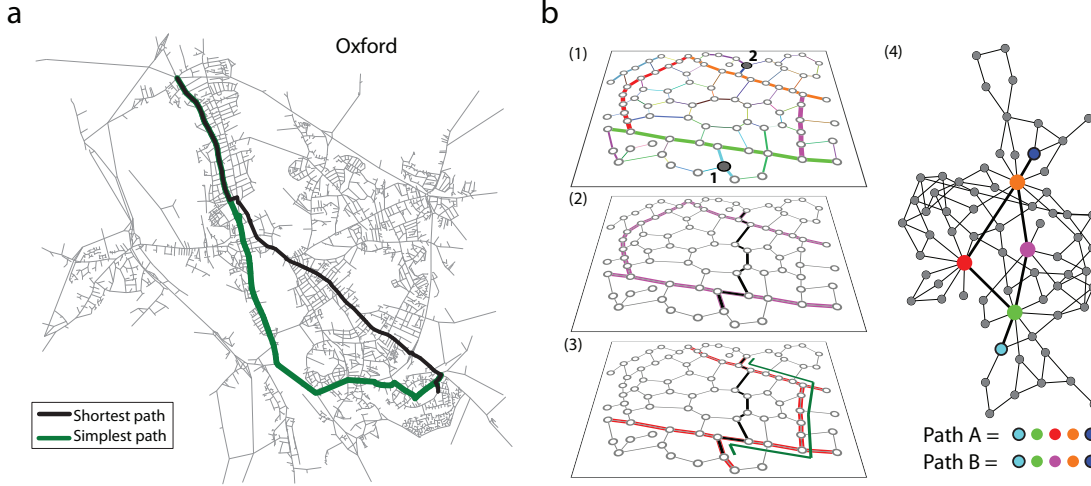


Figure 3.8: **Simplest path: example and calculation.** (a) Example of shortest (black line) and simplest (dark green line) paths illustration on the Oxford (UK) street network. The simplest path has less turns at the expense of being longer than the shortest path. (b) In (1) we show a planar network and in (4) its dual representation. The colours of straight lines in (1) corresponds to the ones of nodes in (4). The simplest path between nodes 1 and 2 is obtained by the shortest path in the dual space (between nodes cyan and blue in this case). There are two paths A and B with length 4 (corresponding to magenta lines in (2) and to red lines in (3)), and the shortest one is chosen as the simplest path (green line in (3)). For comparison we also show the shortest path (black line in (2,3)).

The simplicity index and the simplicity profile

Is possible to define the number of turns τ of a given path as the number of switches from one straight line to another when walking along this path. This quantity is intimately related to the amount of information required to move along the path [118], i.e. the number of turns that are necessary to move between two locations. The average number of turns $\langle \tau \rangle$ versus the number of nodes N indeed displays a small-world type behaviour characterized by a slow logarithmic increase with N as can be seen in Fig 3.9, consistently with previous analysis of the dual network [118, 67]. This feature is thus not very useful to distinguish different networks and shows that the distribution of the number of turns is a very partial information and tells very little about the spatial structure of the simplest paths. In fact, as shown in Fig 3.9, the historical centre of Paris results very similar to a pattern of leaf venations. So a better information is given by the comparison of the lengths of the shortest and the simplest paths given by the ratio $\ell^*(i, j) / \ell(i, j) \geq 1$. It is then natural to introduce the *simplicity index* S as the average

$$S = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{\ell^*(i, j)}{\ell(i, j)} \quad (3.7)$$

The simplicity index is always larger than one and exactly equal to one for a regular square lattice and any tree-like network for example. Large values of S indicate that the simplest paths are on average much longer than the shortest ones, thus that the network is not easily navigable because to reach the destination a great number of turns (informations) are necessary. This new metric is a first indication about the spatial structure of simplest paths but mixes various scales, and in order to obtain a more detailed information, the *simplicity profile* is defined as:

$$S(d) = \frac{1}{N(d)} \sum_{i,j | d_E(i,j)=d} \frac{\ell^*(i,j)}{\ell(i,j)} \quad (3.8)$$

where $d_E(i,j)$ is the euclidean distance between i and j and where $N(d)$ is the number of pairs of nodes at euclidean distance d . This quantity $S(d)$ is larger than one and its variation with d informs us about the large scale structure of these graphs. In order to understand the proposed simplicity profile it is possible to draw a generic shape of it: for small d , at the scale of nearest neighbours, there is a large probability that the simplest and shortest paths have the same length, yielding $S(d \rightarrow 0) \sim 1$, and increasing for small d . For very large d , it is almost always beneficial to take long straight lines, thus reducing the difference between the simplest and the shortest paths. As a result we expect $S(d)$ to decrease when $d \rightarrow d_{max}$. The simplicity profile will then display in general at least one maximum at an intermediate scale d^* for which the length differences between the shortest and the simplest path is maximum. The length d^* thus represents the typical size of domains not crossed by long straight lines. At this intermediate scale, the detour needed to find long straight lines for the simplest paths is very large.

A null model

In order to provide a simple benchmark to further analyse the results obtained by these new metrics a null model is introduced. Let be N points randomly distributed in the plane and construct the Voronoi graph. It is possible then add a tunable number of straight lines of length ℓ distributed according to $P(\ell) \sim \ell^{-\alpha}$. An examples of networks generated by this model is shown in Fig 3.10. This simple null model can be used to understand if the density of straight line is the only parameter that effect the simplicity profile.

Gini coefficient

The Gini coefficient quantifies the inequalities of the lengths of straight lines, and is defined as in [119]

$$G_k = \frac{1}{2E^2\bar{\ell}} \sum_{i,j=1}^E |\ell_i - \ell_j| \quad (3.9)$$

where $\bar{\ell}$ is the average length of straight lines and E is the number of straight lines. The Gini coefficient lies in the range $[0, 1]$ and $G = 0$ when all lengths are equal. On the other hand, if all lengths but one are very small, the Gini coefficient will be close to 1.

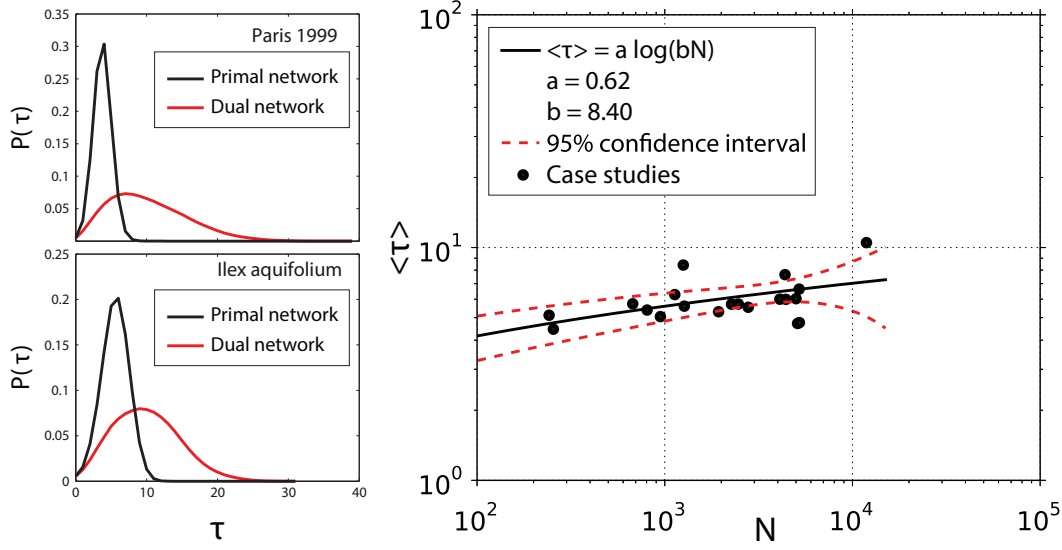


Figure 3.9: **Dual network degree distribution** (a) Probability distribution of number of turns for Paris (1999) and a leaf (*Ilex aquifolium*) for their primal and dual representations. (b) Average number of turns $\langle \tau \rangle$ versus the size of the network N for all the networks studied here. The fit here is logarithmic showing a very slow dependence of τ versus N , a behaviour typical of small-world networks.

3.2.3 Results

The simplicity index S for the various datasets and for the null model as well has been analysed. The results are shown in Fig. 3.13 as a function of the density of straight lines ρ and the Gini coefficient G for the length of straight lines. The density ρ of straight lines is defined as the ratio of total length of straight lines, over the total length of the network, and G is an indicator of the diversity of the length of straight lines. The first observation from Fig. 3.13 is that the simplicity index encodes information which is neither contained in the density ρ nor in the Gini coefficient G , and reveals how the straight lines are distributed in space and participate in the flows on the network.

In Fig. 3.13a, it is possible to observe that the density of straight lines is always larger for urban systems. More precisely, in the biological systems the density lies in the range $\rho \in [0.55, 0.7]$, while we observe $\rho > 0.7$ for artificial systems. Except for the *Physarum*, which appears to be close to a regular lattice with a small simplicity and small Gini coefficient, the simplicity index for the wing and the leaves is larger than the values obtained for the null model. These results indicate that the organization of straight lines in biological systems is very different from artificial systems, that have very similar values of ρ , G , and S . In particular, we observe a hierarchy of straight lines in biological systems (see Fig 3.11): a main artery (the midrib for leaves) connects to veins which in turn are connected to smaller veins and so on. In the case of dragonfly wing, the main straight line is given by the external border of the network. The existence of these main straight lines in biological systems will impact the structure of simplest paths and impose some large detour, resulting in a larger value of the simplicity index.

For urban systems, the simplicity is very close to the null model (of order 1.3 in this density range),

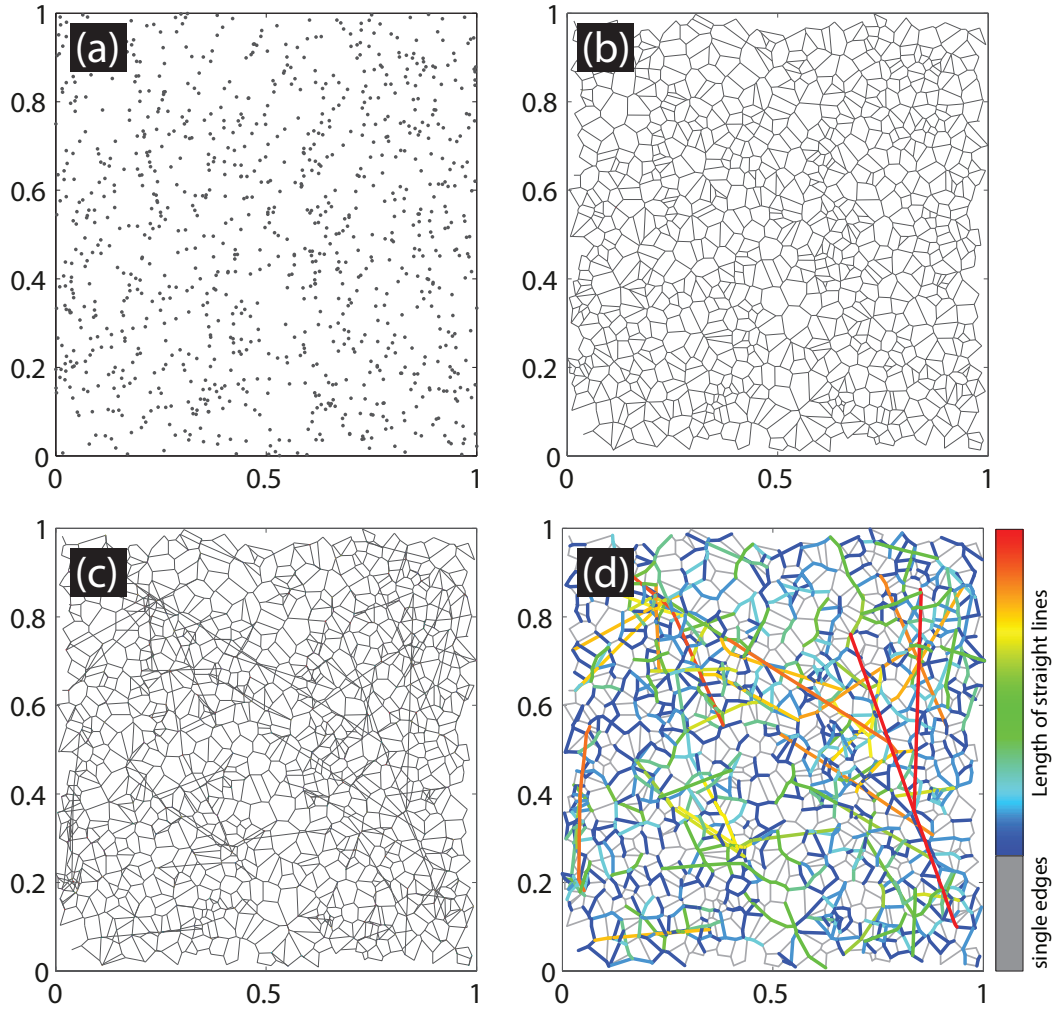


Figure 3.10: **Illustration of the model:** from a random set of points (a), (b) the Voronoi lattice over this set has been constructed and (c) add straight lines of random lengths and random locations. In (d) the straight lines with a colour code according to their length are represented, and single edges are represented in grey.

suggesting that in dense urban systems, long straight lines are added at random (An exception concerns, the pre-Haussmannian Paris (1789-1836) for which we observe a simplicity smaller than for the null model, the reason being probably that the networks at these times were very sparse). As a result, navigation on urban systems requires relatively less information with no additional cost: the simplest path is not too different from the shortest path.

Finally, it is possible to note an interesting effect in the null model in Fig. 3.13a which is the existence of a maximum of the simplicity at densities of order $\rho \sim 0.55$. In this density regime, using straight lines implies having to make large detours. However, when the density exceeds 0.6, there are enough straight lines to enable a simplest path which differs not too much from the shortest one.

We can now discuss the simplicity profiles shown in Fig. 3.11. It is possible to observe that basically, for most of these systems, the simplicity profile displays the generic shape with a maximum at an intermediate scale. In urban cases, such as Bologna and central Nantes, we have a typical mono-centric system with a dense center and a few important radial straight lines, leading to a simple profile $S(d)$. In the case of Oxford and the Australian highway network, the poly-centric organization leads to multiple peaks in the simplicity profile (Fig. 3.11). Interestingly, we observe that the profiles for Australian highways and railways in the UK are very different, despite their similar scale, density ρ , and Gini coefficient G . In particular, the UK railway displays small values of the simplicity (less than $\lesssim 1.2$) while for the Australian highway network there are many pairs of nodes for which the simplest path is much longer than the shortest one. We also observe that the profile for both street and water systems of Nantes have a very similar shape, pointing to the fact that these networks are strongly correlated. In addition, the position and the height of the peak (≈ 1.4) observed for the Nantes water system suggests that this distribution system has similar features compared to biological systems such as vein networks in leaves (see below) whose function is also distribution.

Compared to urban systems, the simplicity profile of biological networks have a single well-defined, and much more pronounced peak. We observe values of order $S_{max} \approx 1.5$ and 2.5 for $d^*/d_{max} \approx 0.2$, meaning that for this range of distance, the detour made by the simplest path is very large. This peak is related to the existence of domains of typical size d^* not crossed by large veins. We see here a clear effect of the existence of the spatial organization of long straight lines in these systems, probably optimized for the distribution (of water for leaves). The decay for large d is also much faster in the biological case compared to urban systems: this shows that in biological systems there are long straight lines allowing to connect far away nodes. This is particularly evident on the leaves shown in Fig. 3.11 where we can see the first levels (primary and secondary) veins, the rest forming a network. For streets, the organization is much less rigid and the hierarchy less strict: we have a more uniform spatial distribution of straight lines, leading to a smoother decrease of $S(d)$.

Going beyond static networks, we apply our new metrics to study the structural changes of time evolving networks.

Urban systems The first example of a time evolving network is the road network of the Groene as studied in the previous section as described in Chapter 2.1 [120]. We have 7 snapshots of this network for different times from 1833 to 2007 (see the Supplementary Information for details). This region evolved without central planning and is thus a good example of an ‘organic’ evolution of urban systems.

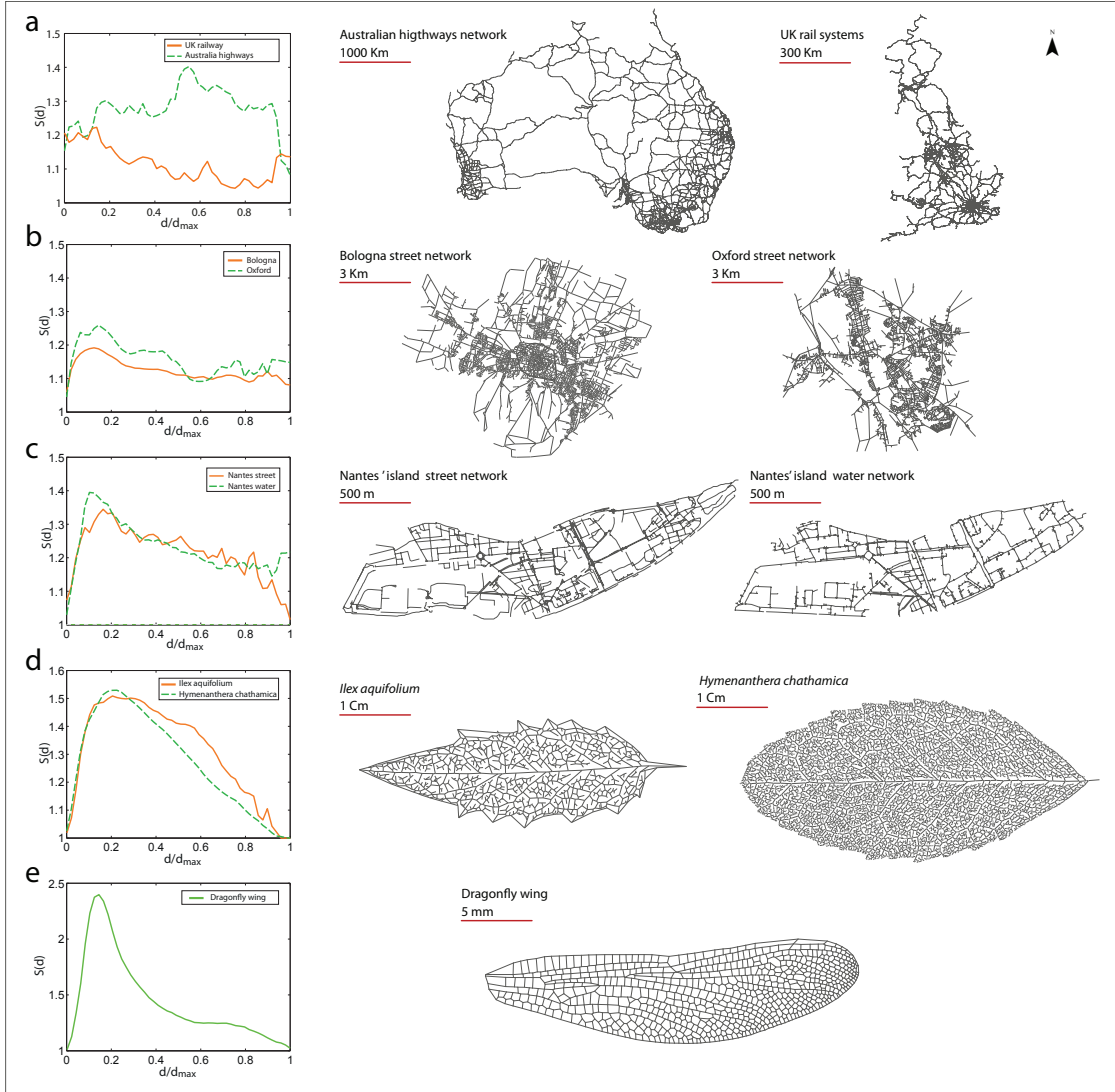


Figure 3.11: **Simplicity profiles.** The simplicity profiles for different networks ranging from large scale networks ($10^6 m$) to small scales of order $10^{-3} m$. Is possible to see on these different examples the effect of the presence of long straight lines and of a poly-centric structure. In particular for cases (d,e), we can clearly see that the peak at $d^* \sim 0.2d_{max}$ corresponds to the size of domains not crossed by long straight lines.

The simplicity profile shown in Fig. 3.12(a) allows us to distinguish two different periods. The first period from 1833 to 1955 displays a relatively small simplicity at all scales, while a distinct second regime appears from 1980 until now. In this latter regime, the simplicity profile is substantially larger for all scales. This is an effect of the massive urban intensification, leading to a poly-centric structure where the readability and the ease to navigate are drastically lowered.

At a smaller scale, we study the evolution of central Paris between 1789 and 1999. This dataset provides an interesting case study, as Paris experienced large changes due to Haussmann in the middle of the 19th century (see [71] and Sec.2.1 for details and more references about this network). This is an opportunity to observe quantitatively the effect of top-down planning: until 1836, we are in the pre-Haussmann Paris, while from 1888 until now we are in the post-Haussmann period. The effect of Haussmann's central planning is clearly visible on the network shown in Fig. 3.12(b). From 1789 to 1836, we have a relatively large simplicity at all scales and we observe a decrease in that period at small scales ($d/d_{max} \lesssim 0.4$) which corresponds well to the fact that many religious and aristocratic domains and properties were sold and divided in order to create new houses and new roads, improving congestion inside Paris. The 1826-1836 transition displays a decrease of the simplicity for distance larger than roughly 5 kms (corresponding to $d/d_{max} \approx 0.6$) indicating that long distance routes were simplified. It is interesting to note that during this period the eastern part of Paris experienced large transformations with the construction of the channel St. Martin. Finally in the period 1836 to 1888, when Paris experienced Haussmann's transformation, the simplicity profile is strongly affected: compared to 1836, the simplicity is improved in the range $d/d_{max} \in [0.3, 0.8]$, which can be attributed to the construction of large avenues connecting important nodes of the city. In addition, we observe the surprising effect that at large scales $d/d_{max} \gtrsim 0.8$, the simplicity is degraded by Haussmann's work: this however could be an artefact of the method and the fact that we considered a portion of Paris only and neglected the effect of surroundings.

Finally, it is possible to note that differences between Groane and Paris might be explained in terms of a sparse, poly-centric urban settlement (Groane) versus a dense one (Paris). In particular, in the 'urban' phase for Groane (after 1955), the simplicity profile becomes similar to the one of a dense urban area such as Paris.

Finally, we show the results in Fig. 3.12(c) for the *Physarum Polycephalum*, a biological system evolving at the centimeter scale. *Physarum* is a unicellular multi-nucleated amoeboid that during its vegetative state takes a complex shape. Its plasmodium viscous body whose goal is to find and connect to food sources, crystallizes in a planar network-like structure of micro tubes [121]. In simple terms, *Physarum*'s foraging strategy can be summarized in two phases: i) the exploration phase in which it grows and reacts to the environment and ii) the crystallization phase in which it connects to food sources with micro-tubes. We inoculated active plasmodium over a single food source and observe the micro-tube network at six phases of its growth. Under these conditions, we observe that the network is statistically isotropic around the food source as shown in Fig. 3.12(c) and develops essentially radially. We first observe that the simplicity profile for the *Physarum* is relatively low (less than $\lesssim 1.2$), suggesting that simplicity could be an important factor in the evolution of this organism. A closer observation shows that during its evolution, the *Physarum* adds new links to the previous network and also modifies the network on a larger scale, as revealed by the changes of the simplicity profile. The evolution of the profile is similar to the one obtained for the null model when the density is increased (see SI),

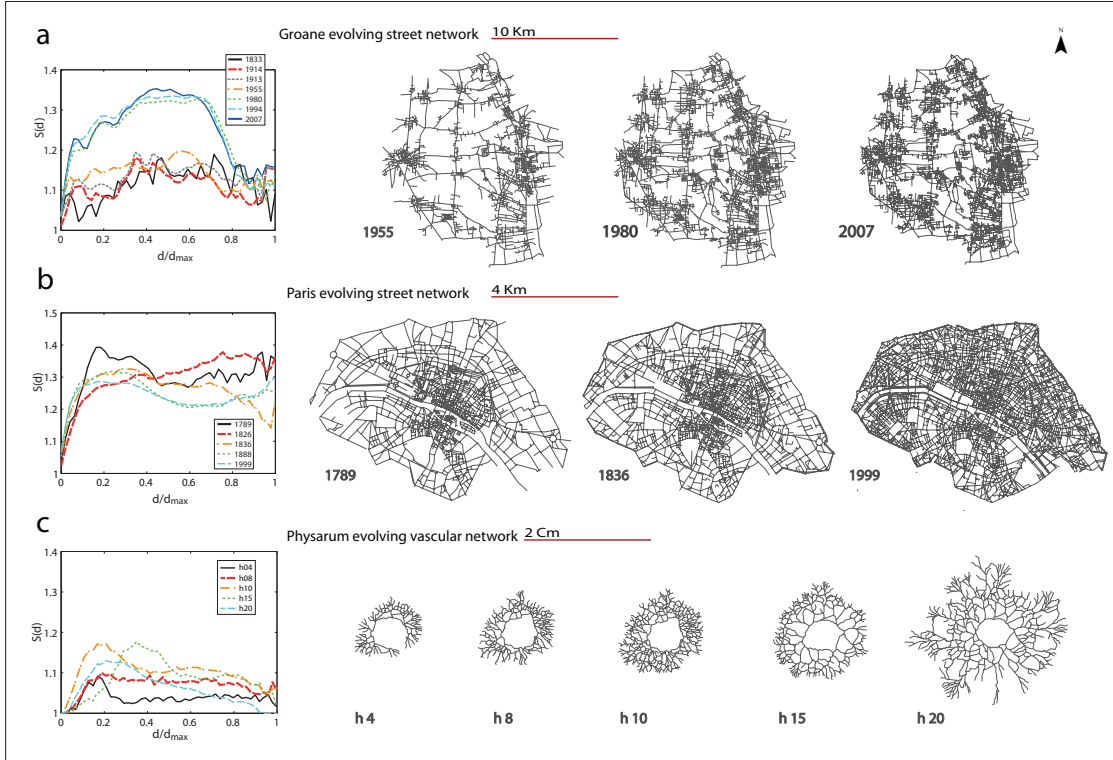


Figure 3.12: **Simplicity profiles for time-varying networks.** Is represented here the profiles for (a) the road network of the Groane region (Italy), (b) the street network of Paris (France) in the pre-Haussmannian (1789, 1836) and post-Haussmannian (1999) periods, and in (c) the *Physarum* network growing on a period of one day approximately. We observe on (a) and (b) that the evolution of the profile is able to reveal important structural changes. In (c) the evolution follows closely the one obtained with the null model.

suggesting that the main growth mechanism could be described as essentially a random addition of straight lines.

3.2.4 Discussion

The new metrics introduced in this section encode in a useful way both topological and geometrical information about the structure of planar graphs at different scales. In particular, the results highlight the structural differences between biological and artificial networks. In the former, we have a clear spatial organization of straight lines, with a clear hierarchy of lines (midrib, veins, etc), leading to simplest paths that require a very small number of turns but at the cost of large detours. In contrast, there is no such strong spatial organization in urban systems, where the simplicity is usually smaller and comparable to a null model with straight lines of random length and location. These differences between biological and urban systems might be related to the different functions of these networks: biological networks are mainly distribution networks serving the purpose of providing important fluids and materials. In contrast, the role of road networks is not only to distribute goods but to enable individuals to move from one point of the city to another. In addition, while biological networks are

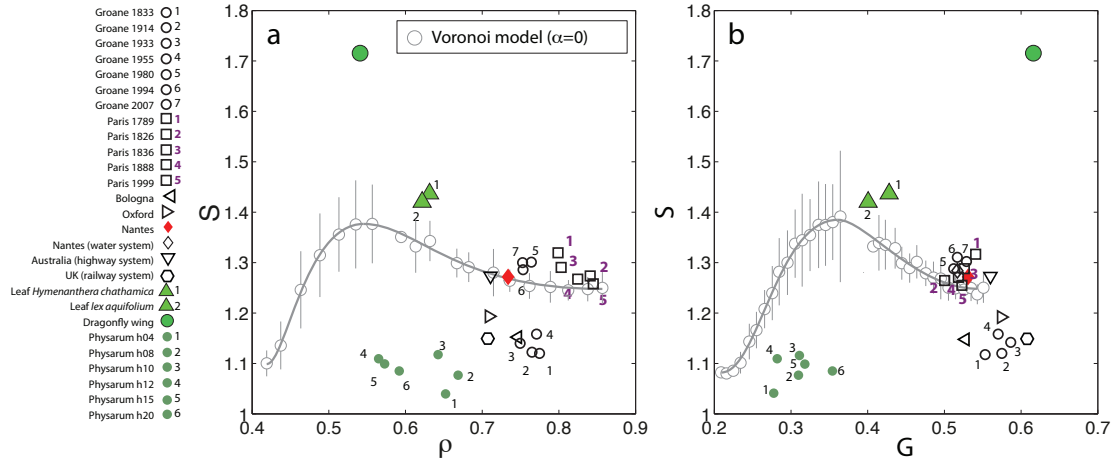


Figure 3.13: **Simplicity index.** Simplicity versus (a) the density of straight line ρ and (b) the Gini coefficient for the length of straight lines. In both plots, the symbols correspond to the different networks studied here. We also represented the result for the null model (for $\alpha = 0$) and its cubic spline interpolation (continuous line). From (a) we see that biological networks are limited to the region $\rho \leq 0.7$ and have a large simplicity index, and from (b) we see that urban networks have simultaneously higher values of G and relatively small values of S .

usually the result of a single process, urban systems are the product of a more complex evolution corresponding to different needs and technologies. These new metrics also allow us to track important structural changes of these networks. The simplicity profile thus appears as a useful tool which could provide a quantitative classification of planar graphs and could help in constructing a typology of leaves or street patterns.

3.3 Multiplex approach to urban network analysis

The following study is focused on multiplex spatial networks in urban environment. Results of the study has been published in the following scientific paper.

Strano E, Shay S, Dobson S, Barthélemy M, (2015) *Multiplex networks in metropolitan areas: generic features and local effects*. Royal Society Interface, 12 (111), 20150651.

The candidate contributed to the study in following phases: design of the study, preparation of the data, analysis of the results and writing of the the paper. Data analysis have been made together with Sara Shay, the second author of the paper. This paper has been reviewed by several media such as WIRED, Scientific American and BBC news. The study made the cover of Royal Society Interface 12(112).

3.3.1 Introduction

As described in Chapter 1.3 urban road networks have been largely studied by means of complex spatial networks. In parallel, there are also studies on the structure of subway networks [122, 123, 124], their evolution [125], and their robustness [126, 72]. However, these networks are *not* independent, and the important result in [127] showed that coupling between networks can be critical and can affect the global behaviour of a system. It is in this context that multilayer (or multiplex) networks [128, 129, 130, 131] are studied and provide the convenient conceptual framework. A few recent studies considered the impact of the multilayer structure on various general processes [132, 133], and specifically in the case of transportation networks [134, 135]. Multilayer networks offer a good theoretical framework for understanding how interconnected transportation networks are shaping cities and how they may affect their operation. Moreover, given the increasing interest in urban systems, an empirical study of the impact of the multilayer structure of transport systems on mobility appears both crucial and timely.

In this section the mutually connected underground and street networks in the large metropolitan areas of Greater London and New York City have been studied with the aim to understand their coupling affects their global properties. In particular, the effect of varying the subway speed is analyzed. It is shown that increasing the underground speed can lead to unexpected counter-effects. The analysis focuses on three main network features and findings: i) the behavior of quickest paths at the city scale; ii) the definition and study of two new measures called local outreach and the urban horizon; and iii) the effect on the spatial distribution of betweenness centrality given by intermodality. It is important to stress that studies of urban transportation networks have important implications for urban policies and private investment, and in general play an important role in the urban planning chain. In fact, inter-modality transportation efficiency and simulations have been extensively studied in the transportation engineering literature [136], where the typical supply-demand approach prevails but where the analysis of topological properties of networks is almost wholly neglected and where the different transportation modes are often treated separately. One goal in this study is to shift the focus only topological coupling aspect of transportation network design: we show this to be extremely relevant, and suggests that the multilayer network view of these systems should be integrated into elaborated models of urban and transportation planning.

3.3.2 Material and Methods

Using data from Open Street Map [89], we construct both the street and the subway networks for London (UK) and New York City (USA). We downloaded data on street networks and underground in geo-referenced vectorial format from Open Street Map, which contains detailed streets and rail tracks networks, including train depots and double tracks. (The rationale behind the geographical extent of these networks is to include the full underground systems and surrounding street networks.) In addition, a series of automatic and manual topological cleaning operations were needed in order to extract consistent and usable graphs as described in Sec.3.1. The size and geography of the two cities is clearly different as we can observe it in Fig. 3.14a,b.

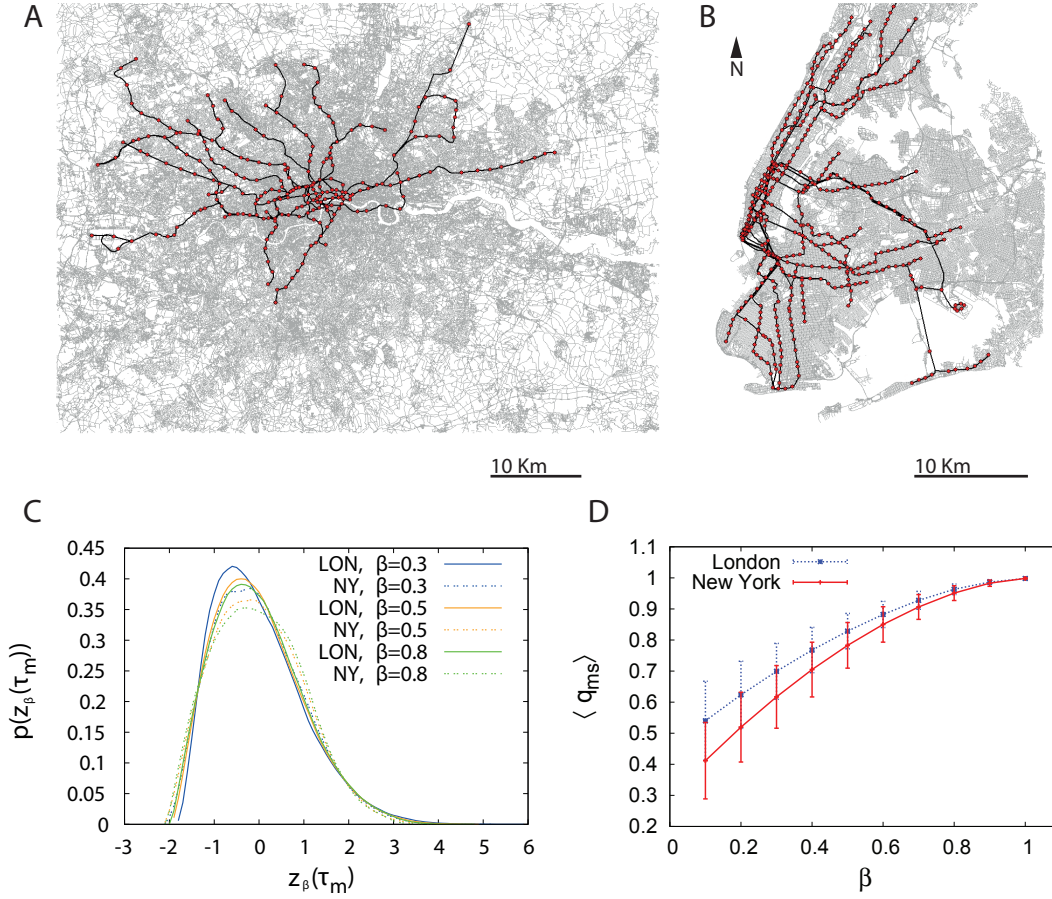


Figure 3.14: **The spatial extent of the two metropolitan areas.** Note that the Greater London area is not entirely covered by the underground system, in contrast with New York where most areas are connected by the subway. (C) Distribution of normalised quickest path times (computed in the multilayer system) $z_\beta(\tau_m) = \frac{\tau_m - \langle \tau_m \rangle}{\sqrt{\text{Var}(\tau_m)}}$. (D) The quantity $\langle q_{ms} \rangle$ averaged over all nodes as a function of β (the error bars indicate here the dispersion around the average). The average ratio of travel times with and without the subway layer is typically of order 0.5 and does not vary much with β .

We can thus obtained the weighted graph $G_s = (V_s, E_s)$ of the connected street network in its “primal” representation, with nodes being street junctions and edges representing the street segments connecting them, and the weights given by the street length $W = \text{streetlength}(m)$. Similarly, we obtained the connected underground network $G_u = (V_u, E_u)$ with nodes representing underground stations and links connecting successive stations on the same line, and weighted by the length of the line segment $W = \text{linklength}(m)$. From a theoretical point of view, the interdependent or multilayer [128] network, G_{multi} is defined as the union of these two networks. Here we have subway stations and road intersections that we consider to be different nodes. Underground stations are accessible from more than one access on the street, but for the sake of simplicity we construct the multilayer network by connecting each underground station to its closest street junction only (a simplification that won't change the structure of quickest paths).

3.3.3 Results

The generic nature of quickest paths

New York is composed of two large and almost disconnected components with the underground systems covering a similar spatial extent. London instead presents – at a large scale – a typical radiocentric urban structure with the underground systems connecting satellite districts and peripheries to the urban core. Differences both in size and geography between these cities are also reflected by basic network descriptors shown in Table 3.1.

Case	N	M	Total length (km)	\bar{l}_{ij}^{geo} (km)	\bar{l}_{ij}^{topo}	$Diam^{geo}$ (km)	$Diam^{topo}$	$\sqrt{A}(\text{km})$
London streets	324536	427920	34493.73	25.83	178.16	89.31	368	
London tube	263	296	385.98	18.55	14.26	60.3	42	-
London multi	324799	428479	34886.52	25.78	96.16	89.27	288	-
New York street	68417	112827	12153.81	17.94	106.64	55.18	278	34.93
New York subway	454	489	416.12	18.87	19.10	57.28	62	-
New York multi	68871	113770	12579.44	17.91	54.45	55.18	205	-

Table 3.1: **Basic features of multiplex urban network.** The number of nodes N , the number of links M , the cost defined as the total length of all links, the average geographical and topological shortest path length, \bar{l}_{ij}^{geo} (km) and \bar{l}_{ij}^{topo} respectively; the maximum geographical and topological shortest path length, D^{geo} (km) and D^{topo} respectively; the density of streets $\rho_s = \frac{N_s}{A}$ and the density of underground stations $\rho_u = \frac{N_u}{A}$, where A is the area of each city.

For both cities, the (spatial) diameter of the multiplex is essentially dominated by the street network. We can also observe that the topological diameter of the multiplex is lower than the street layer, thanks to the subway structure allowing for topologically shorter paths. The efficiency of the subway is however also due to its speed which is in general larger than that of overground modes such as private cars, taxis, or buses. In order to reflect this, a parameter $0 < \beta \leq 1$ is introduced to describe the ratio of the speeds between the two systems, similarly to the theoretical analysis proposed in [134]. This parameter β measures the travel cost in time units associated to the underground links. This means that the number of time units taken to traverse an underground link of length l meters is βl , which is $\frac{1}{\beta}$ times faster than the time taken to traverse the same length on the street network. Thus a smaller β corresponds to

a faster underground speed, as compared with the speed on the street network. The introduction of this parameter allows us to study the properties of the multilayer system as a function of underground speed. Naively one could expect that the system as a whole will be more efficient for faster subways, but it is shown here that it is not always the case. Finally β can be measured empirically, that is for London $\beta_{London} \approx 0.48$ and a slightly larger value for NY $\beta_{NYC} \approx 0.55$.

Let denote by $\tau_s(i, j)$ the travel cost (*i.e.*, the number of time units) of the quickest path between street nodes $i, j \in V_s$, and by $\tau_m(i, j)$ the cost of the quickest path between i and j in the multilayers network (*i.e.*, a path which can traverse *both* street *and* underground links). The normalised quantity

$$z_\beta(\tau_m) = \frac{\tau_m - \langle \tau_m \rangle}{\sqrt{Var(\tau_m)}} \quad (3.10)$$

displays a behaviour, in temr of functional form, that is roughly constant for β larger than 0.2 – 0.3, as shown in Fig. 3.14c, demonstrating that the effect of β is essentially contained in the average and variance of τ_m . This is a rather surprising result, given that the two cities display many geographical and structural differences.

The cost $\tau_m(i, j)$ between nodes i and j can be written as

$$\tau_m(i, j) = \sum_{e \in P(i, j)} \tau(e) \quad (3.11)$$

where the sum is over all links e that belong to the quickest path $P(i, j)$ and where $\tau(e)$ is the cost on this link. (We neglect inter-modal change costs in this simple argument.) If the path is long enough, and if the random variables $\tau(e)$ do not display long-range correlations and are not broadly distributed, the central limit theorem applies and the distribution of the τ_m follows a Gaussian distribution in a certain range. There are obviously deviations observed for small values of β coming from the fact that the paths' durations become very heterogeneous depending on the proximity of their origin or destination to subway stations. In this respect, a very high relative subway velocity enhances spatial differences in the city and may lead to an uneven distribution of accessibility, a fact that will be confirmed below with the local outreach analysis.

The average ratio between the travel costs from i to other street nodes through the multilayer network and through the street network can be defined as

$$q_{ms}(i) = \frac{1}{N_s - 1} \sum_{j \in V_s} \frac{\tau_m(i, j)}{\tau_s(i, j)} \quad (3.12)$$

where N_s is the number of street nodes. The larger this ratio, the larger the effect of the underground on travel costs. We see in Fig. 3.14d that typical values are of order 0.5 for both cities and that the effect of

β is rather weak: a decrease from $\beta = 1$ to $\beta = 0.5$ leads to a decrease of $\langle q_{ms} \rangle$ of order 20%. In addition, it seems that the effect of subways in London is less important than in New York, which is probably due to the lesser extent of the subway in the Greater London area.

A central quantity for describing the importance of inter-modality is given by

$$\lambda(i, j) = \frac{\sigma_{i,j}^{\text{multi}}}{\sigma_{i,j}} \quad (3.13)$$

where $\sigma_{i,j}$ is the total number of shortest paths between i and j (using either one or two networks), and $\sigma_{i,j}^{\text{multi}}$ the number of paths using edges of both networks at least once. It characterizes the importance of multi-modality for the path from i to j . If we sum over all possible destination nodes j , we can quantify the added value of the interlayer coupling to the reachability of nodes, and obtain the interdependency [134] of a street node $i \in V_s$ defined as

$$\lambda(i) = \frac{1}{N_s} \sum_{j \in V_s} \frac{\sigma_{i,j}^{\text{multi}}}{\sigma_{i,j}} \quad (3.14)$$

(Note that a similar measure has been used in the transportation design literature under the name of *inter-modal connectivity* [136].) In order to understand the effect of scale on the interdependence, the interdependence profile is defined as:

$$Q_\lambda(d) = \frac{1}{N(d)} \sum_{\substack{i,j \in V_s \\ d_e(i,j)=d}} \lambda(i, j) \quad (3.15)$$

where $d_e(i, j)$ is the Euclidean distance between i and j and $N(d)$ is the number of pairs of nodes at Euclidean distance d . In Fig. 3.15a is shown the average interdependence among all street nodes as a function of β and the resulting interdependence profile Fig. 3.15b.

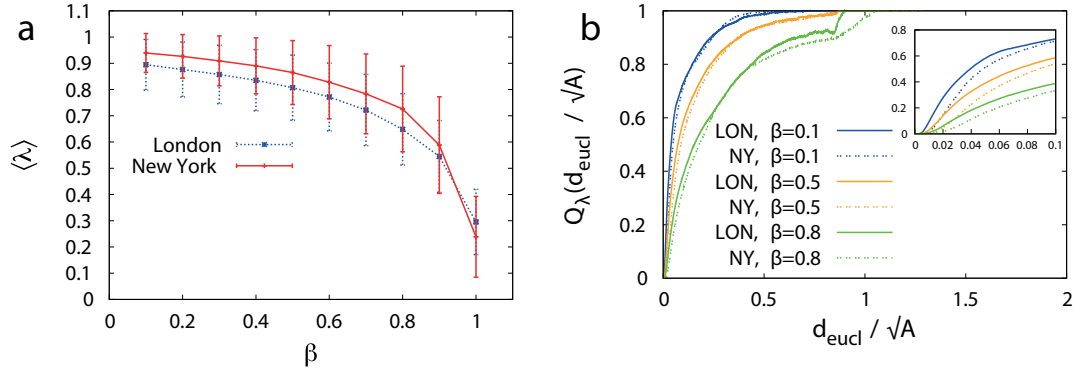


Figure 3.15: **Average interdependence.** (a) Average interdependence, $\langle \lambda \rangle$ as a function of β . (b) Normalised interdependence profile computed for different values of β . Both cities exhibit a similar behaviour despite very different geographical structures.

We see from these figures that, in both cities, the existence of the underground has a very large impact. For example, for $\beta = 0.8$ we obtain λ around 0.7, meaning that even when the underground is only 1.25 times faster than the street network, already about 70% percent of the quickest paths are going through the underground. A slight decrease of β for β close to one thus has a large impact on structure of quickest paths, while for smaller values of β , improving the subway speed does not bring a significant improvement of the structure of quickest paths. In both cities there is a sharp increase in λ for small Euclidian distance, meaning that already for relatively short trips, it is worth “hopping on” to the underground. (Note that we neglect here waiting, walking, and connecting times which can be significant [135].) The slope of the interdependence profile at small $d_{\text{eucl}} \approx 0$ is increasing as β is decreasing, suggesting that a slight increase in the underground speed could make the networks highly interdependent even at very small scales.

Both cities therefore display a remarkably similar behaviour over all these interdependency-related quantities (in particular, see Fig. 3.15(b)), suggesting here again a possible common behaviour for multiplex transportation network in cities. While further studies are needed in order to substantiate a claim of “universality”, our results point to the possible existence of some kind of a statistical law of large numbers that applies to quickest paths in multiplex urban transportation networks.

We note that it is not trivial that the central limit theorem applies here and it doesn’t mean that the network topology is irrelevant. The fact that we can sum a large number of quantities, which are uncorrelated (a necessary condition for the central limit theorem to apply) comes from the specific structure of these transportation systems (spatial constraints for example certainly play an important role). In addition, more complex quantities (such as the interdependence for example) also display a large level of similarity for the two cities, a fact that cannot at this stage be simply related to a central limit theorem. These different results point to the potentially useful fact that actually few parameters seem to govern the behavior of these quantities, which could lead to many useful simplifications in more elaborated models that contain a large number of parameters.

Local outreach and the urban spatial horizon

The presence of a transportation mode such as a subway affects the overall performance of a city in terms of efficiency of transport and the accessibility of certain locations, but also has an important impact on how pairs of locations are connected. In order to measure this effect, we define the *spatial outreach* of a street node $i \in V_{\text{street}}$ as the average Euclidean distance from i to all other street nodes that are reachable within a given travel cost, τ :

$$L_\tau(i) = \frac{1}{N(\tau)} \sum_{j|\tau_m(i,j) < \tau} d_e(i, j) \quad (3.16)$$

where $d_e(i, j)$ is the Euclidean distance between node i and j , and $N(\tau)$ is the number of nodes reachable on the multilayer network within a given travel cost τ . In Fig. 3.16 we show the average local outreach as a function of the travel cost threshold τ , which displays a non-linear behaviour due to the different speeds achievable in the two transportation modes. This provides support for a general effect already known: for longer trips, faster transportation modes are used (see for example [135] for the UK case).

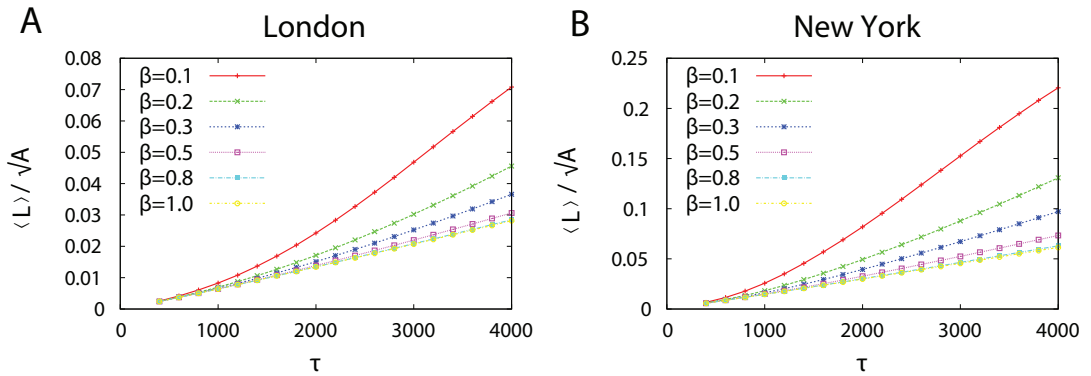


Figure 3.16: **Local outreach.** Average local outreach $\langle L \rangle$ normalised by the square root of area, for London (a) and New York (b).

As shown in Fig. 3.17, a, b, d, e as β decreases, the nodes having a high local outreach are concentrated close to underground stations where the underground is the most accessible, and the graph consisting of high-outreach nodes (red nodes on the map) becomes less fragmented.

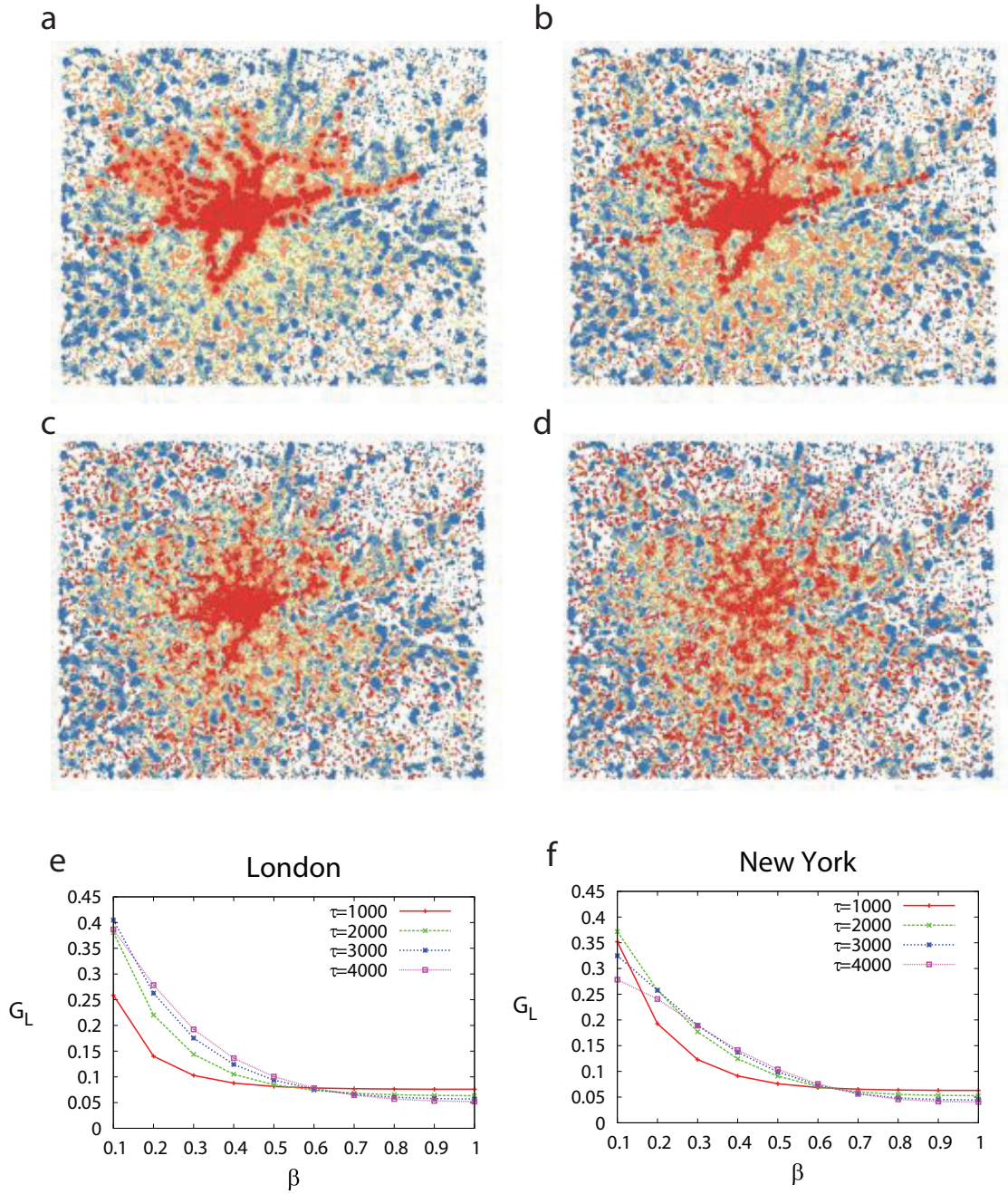


Figure 3.17: **Spatial distribution of the local outreach in London.** Maps showing for London the spatial distribution of the local outreach defined in Eq. 3.16 for a travel cost $\tau = 2000$ and speed ratio (a) $\beta=0.1$, (b) $\beta=0.4$, (c) $\beta=0.7$ and (d) $\beta=1$. As the underground's speed is increasing compared to the speed of the street network (*i.e.*, decreasing β) the nodes having a high local outreach are concentrated at the center along the underground network, where the underground is most accessible. These figures were created using ESRI ArcMap 10.1 and Adobe Illustrator. (e, f) The graphs show the Gini coefficient of the local outreach versus β for both cities.

In other words, as the underground becomes faster, a continuous area of high-outreach nodes emerges (the *commutable zone*) in the city centre and around the nodes of the underground network, implying that a person can travel from this area to faraway places (large Euclidean distance) at a small travel cost τ . The location of this highly accessible zone cluster from a dispersed configuration, as in Fig. 3.17d, to a centralised one, as in Fig. 3.17a, which shows a centralisation effect due to the accessibility provided by the underground. The dispersion of the local outreach also displays a very interesting result demonstrated by its Gini coefficient $G_L \in [0, 1]$. Indeed, in Fig. 3.17e,f we see that for both cities for $\beta > 0.5$ the accessibility is distributed almost uniformly amongst all the places in the cities, while for smaller β (faster underground) the shift to an uneven distribution of accessibility is clear. This result suggests that transportation policies that focus on increasing the speed on a single travel modality may lead to the undesirable spatial heterogeneity in the accessibility of different locations.

We show in Fig. 3.18 the probability that the outreach is larger than a certain fraction αL of the size of the city, and we observe the existence of a threshold α_c .

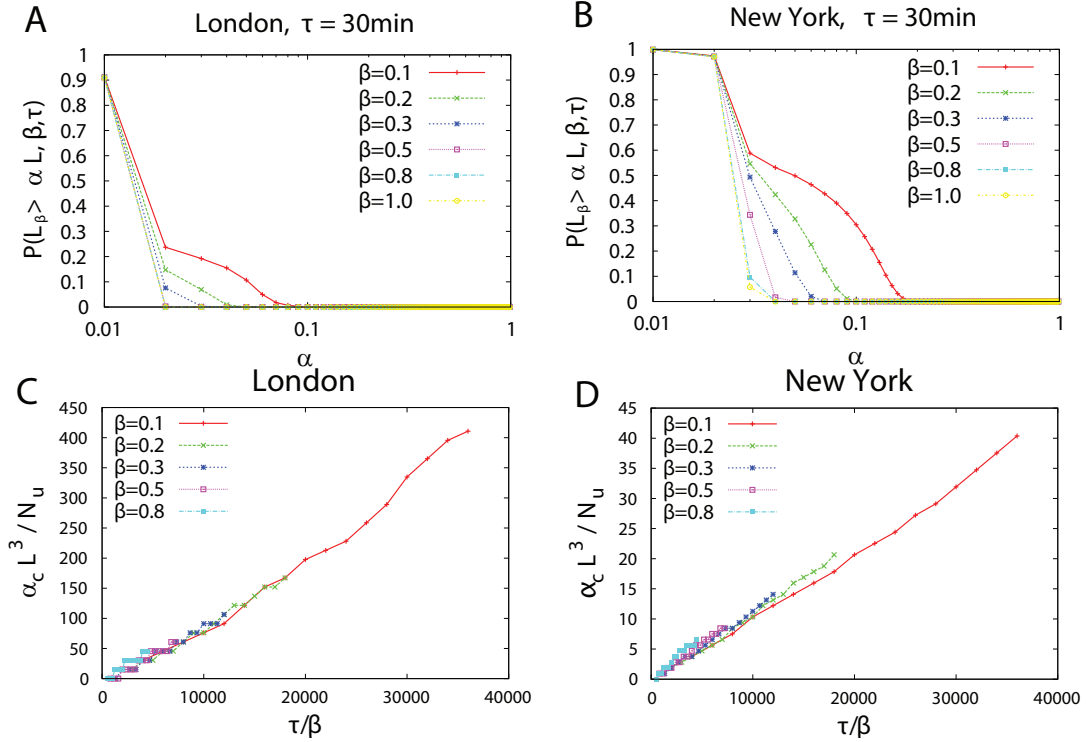


Figure 3.18: **Local outreach** (a)-(b) Fraction of nodes with a local outreach larger than $\alpha\sqrt{A}$ for London (a) and New York (b). (c)-(d) when both axis are rescaled according to Eq. 3.19, we observed a data collapse with all curves collapsing on a single one.

The existence of a threshold less than one means that, for given values of β and τ , there is a maximal value L_m for the outreach. We can estimate the value of L_m by using a simple argument: the maximum value is reached when the path is “essentially” made on the quickest transportation mode, the subway. This transportation mode has a velocity given by v/β , and the probability that a station is within reach

Chapter 3. Results

(in circle of radius d_0 corresponding to the typical walking distance to reach the subway) is

$$p = \rho_u \pi d_0^2 \quad (3.17)$$

where $\rho_u = N_u / A$ is the density of subway station ($A = L^2$ is the area of the city and N_u is the number of subway stations). The maximal outreach L_m is then given by

$$L_m = \frac{N_u}{L^2} \pi d_0^2 \frac{v}{\beta} \tau \quad (3.18)$$

and $\alpha_c = L_m / L$ is then given by

$$\alpha_c = \frac{N_u}{L^3} \pi d_0^2 v \frac{\tau}{\beta} \quad (3.19)$$

This last equation shows in particular that the quantity $\alpha_c L^3 / N_u$ should increase linearly with τ / β , with a constant of proportionality depending on geometry the city, and we observe that this scaling in is agreement with simulations (see figures 3.18c, d). In particular, we see that the slopes for London and New York are different: the ratio of the constant pre-factors is about 10, suggesting that the subway system in London is more efficient in terms of the outreach that can used as a measure of the “urban horizon”.

The geography and distribution of urban centrality

Betweenness centrality (BC) [111] is one of the important quantities in complex networks, as described in Sec. 2.3. In the case of car traffic and congestion, the absence of detailed traffic models or mobility data leads us to use the BC in order to identify the *potentially* congested locations and the effects of spatial structure on the shortest path structure. Even if we know that the assumptions used in the BC calculation can lead to some inaccuracies [137], it is the simplest proxy that contains some level of information about real traffic. We thus explore in this section the spatial distribution of BC in the street network and how it is affected by the underground system. The BC of a street node $v \in V_s$ in the street network is defined as

$$bc_s(v) = \frac{1}{(N_s - 1)(N_s - 2)} \sum_{i,j \in V_s} \frac{\sigma_{i,j}^{\text{street}}(v)}{\sigma_{i,j}^{\text{street}}} \quad (3.20)$$

where $\sigma_{i,j}^{\text{street}}$ is the number of quickest paths between i and j in the street network, of which $\sigma_{i,j}^{\text{street}}(v)$ goes through street node v . Similarly, we define the betweenness centrality of a street node $v \in V_s$ in

the multilayer network as

$$bc_m(v) = \frac{1}{(N_s - 1)(N_s - 2)} \sum_{i,j \in V_s} \frac{\sigma_{i,j}^{\text{multi}}(v)}{\sigma_{i,j}^{\text{multi}}} \quad (3.21)$$

where $\sigma_{i,j}^{\text{multi}}$ is the number of quickest paths between i and j in the multilayer network, of them $\sigma_{i,j}^{\text{multi}}(v)$ goes through street node v .

We can then observe how the parameter β impacts the mobility distribution and the geography of potentially congested areas. The maps in Fig. 3.19(a,b,c,d) show the BC spatial distribution for both cities computed on streets for $\beta = 1$ (a,b) and $\beta = 0.1$ (b,c).

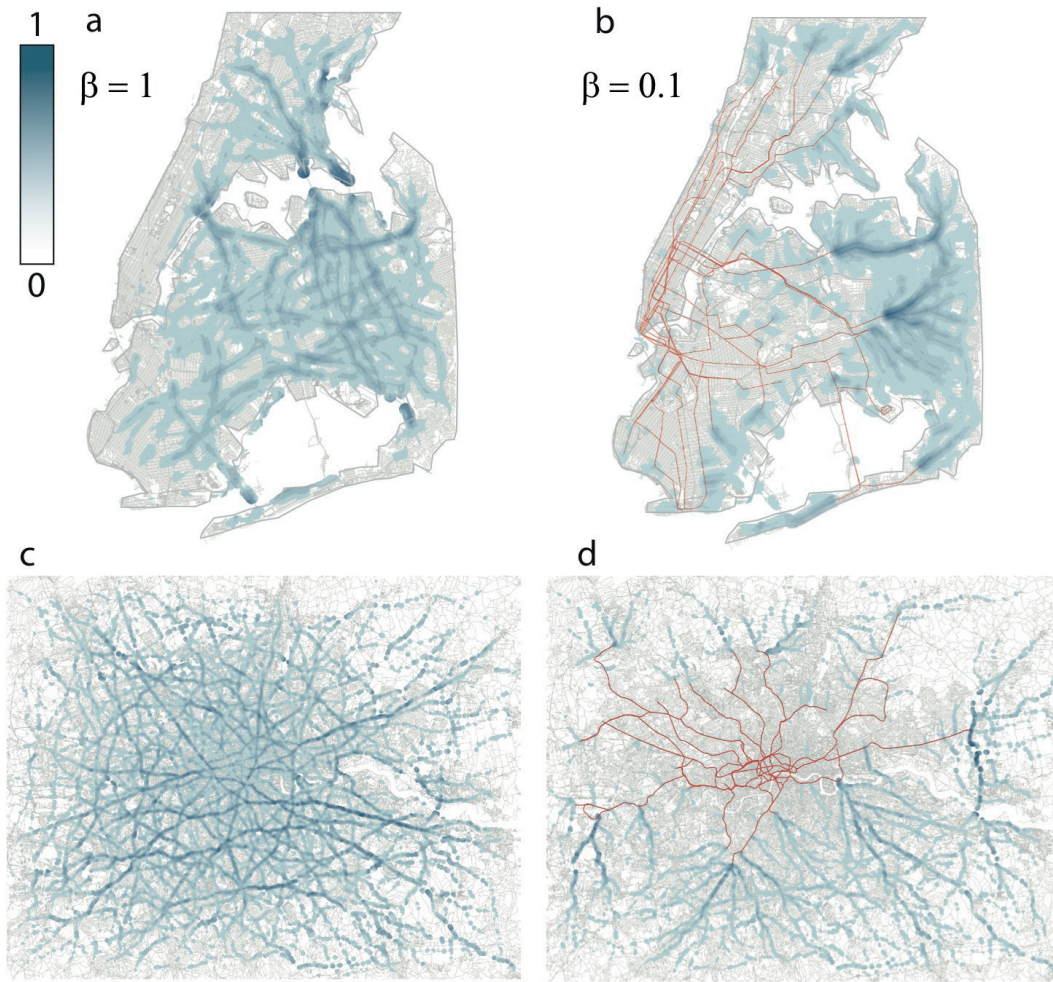


Figure 3.19: **Spatial distribution of BC.** The spatial distribution of BC on the New York (top) and London (bottom) street network for different values of β . We observe a clear crossover from congested road locations for $\beta = 1$ to “focal points” of the underground system for small β (The maps have been obtained using the classical interpolation method Inverse Distance Weigh (IDW) on the street junctions with $Z = BC$). This figure was created using ESRI ArchMap 10.1 and Adobe Illustrator.

These maps clearly display a dramatic change in the spatial distribution of central places when introducing an underground system, shifting congestion from internal street routes and bridges to inter-modal places located at the terminal points of the underground networks, which presumably are used as entry/exit gates for suburban flows to reach core urban areas. Remarkably enough, in both cities, these places are located in urban areas that do not overlap with the underground system, thus possibly creating congestion in unexpected places. In other words, the introduction of underground networks operate as a decentralising force creating congestion in places located at the ends of underground lines and not, for example, in the city centre as one might expect referring to classical results on rewiring processes for chain or lattice networks [70] in which BC is correlated to the distance to the gravitational centre.

3.3.4 Discussion

In this section the effect of underground system on road network metrics has been studied. For quantities relating to quickest paths (interdependency, average quickest path duration), a remarkable similarity emerged between the two considered cities, suggesting the possibility of a universal behaviour. This universality might originate in the fact that the quickest path can be seen as a sum of random variables, which inevitably leads to the central limit theorem. This seems to be the case for the probability distribution of the quickest path time duration, that displays an universal function for a reasonable range of subway speeds. More involved quantities such as the local outreach and the urban horizon also display a simple common behaviour that cannot be recovered using a back-of-the-envelope argument about the quickest path. This possible universality suggests that few parameters govern the behaviour of quickest paths, which could lead to a useful simplifications of elaborated models that contain a large number of parameters. More data on more cities are however needed in order to validate this hypothesis. The spatial distribution of the betweenness centrality can be understood in the framework of another study showing how decentralising housing in London leaves room for commercial activities [138]. Even if the direct causal link between land-use change and transportation network evolution is still not clear [139], presented results seem to confirm this hypothesis. It would be interesting to confront presented results with realistic models used by the transportation community, in particular when the subway speed is modified. More generally, more empirical studies are however needed in order to understand the complex coupling between land-use and the structure of multimodal networks.

It thus seems to be clear that it is important to consider full multimodal, multi-layer network aspects in order to understand the behaviour of an urban transport system – and thus understand the effects of transport on other features of interest. Even if these study is still very theoretical, it shows convincingly that reasoning with only one transportation mode can be extremely misleading, and that policy-makers cannot limit themselves to a single aspect of an urban system without risking making decisions that are locally correct but globally wrong.

Regarding urbanization, a possible interpretation of the results is on the implications of multi-modality on urban growth. Precedent analyses and studies on roads evolution, as presented in Chapter 3.1, have barely considered the effect of a multiplex structure of transportation network. However, specially in the case of early metropolises, such as London and New York, underground evolution had a prominent role in the evolution of suburb neighbourhoods. Presented results suggest that considering the co-evolutionary mutual dependences of roads and other transportation, might be of fundamental importance to model urban growth.

3.4 Scaling of urban economic performance over time and its relationship with the urban land-use

The study focuses on scaling relationships between population and gross metropolitan product in a set of 300 European cities. Part of the presented results has been published as scientific publication to PlosOne:

Strano E, Sood V, *Urban scaling and the European economic transition*. PlosOne (in print).

The candidate conceived and designed the study, prepared and analysed the data, produced the images and wrote the paper. The analyses, in particular the modelling section, have been done together with Vishal Sood, the second author of the paper.

3.4.1 Introduction

As reported in Chapter 1.2 there is now a general scientific consensus on the scaling nature of cities, however its universality is still under debate. In this Chapter the scaling of gross metropolitan product vs population is examined. As recently suggested by Glaeser et al. and Gollin et al. [21, 35] global economy drastically affects urbanization rates and thus it might affect GMP scaling too. In this sense, the assumption that the economic growth of a city is determined solely by its citizens' interactions within urban limits [45] appears questionable. Some natural questions arise when cities are not considered as closed and independent systems. How to explain the consistent poor condition of a metropolis like Lagos? What is the meaning of urban scaling in fast and emerging economies? Is it true that all cities, regardless of their economic environment and their history, are a variation of a theme? Such questions, using archive data on per-capita-gdp and population in 248 cities in the European Union between 2005 and 2010, have been explored. We failed to verify the general consensus that GMP is universally super-linear with population size. European metropolitan areas do not show a uniform and super-linear behaviour in terms of GMP/Population scaling. Former Eastern block cities behave differently than rest of the cities in Europe, suggesting that positive externalities are not sufficient to determine economic growth. We found however that different scaling regimes mirror economic transition and convergence between post-communist countries and former European countries. Results suggest that super-scaling of GMP represents a transition condition of growing economies rather than an universal and stable condition of all cities. Moreover, economic growth of a town is not dependent solely on its population size, but also on higher level economic and social dynamics. GMP in relation to urban land uses and transportation surface has been also explored. The dual behaviour of GMP offers a good case to test the reciprocal interdependences between land-use organization and productivity of a city. Contrary to what expected, the West/East dual behaviour founded in the GMP vs population scaling is not replied by the land-use vs population scaling. It suggests that a fundamental homogeneity in urban spatial organization exists and that it has no effect on urban economic scaling. These findings are important for two reasons: i) they pose urban scaling in a historic perspective, demonstrating how scaling is volatile even in a short time period. ii) they propose a time analysis of urban scaling in fast growing economy demonstrating that productivity growing is independent of population growth and of the land use but rather depend on higher level economic dynamics.

3.4.2 Material and methods

Because of the intrinsic spatial complexity of cities given by their fractal like nature, and because of urban spatial contiguity given by conurbation phenomena, debates and theories around the spatial and political definition of cities as units of analysis are controversial [140] and different approaches have been proposed to solve this problem [61, 60]. Definition of city boundaries is greatly crucial because it defines the unit scale of analysis. This is why it is worth to remember that super-scaling of GMP vs population has been proposed mostly for United States' Metropolitan Statistical Areas that are basically large urbanized areas [141] within the same labour market basin, typically defined by a large and dense urban core including small satellite cities and conurbation land. In this study the equivalent of MSA in Europe, the Large Urban Zone (LUZ) defined by the Statistic European Office, have been used. This allowed to make a comparable analysis with previous literature and to use data on metropolitan population and Gross Metropolitan Product (GMP) over the decade 2005-2010 provided by Eurostats agency [142] for 243 metropolitan area in Europe. In addition to population and GMP data Euro-atlas [93] urban land use data has been used, as described in Chapter 2.2. From Euro-atlas data urban land-use in contrast to agricultural and semi-natural land use have been classified. Fig. 3.20 shows this division. For spatial data there aren't time data. Visual inspection of map of LUZ areas and the metropolitan area provided by EEA confirm that they do overlap in space. A quantitative comparison of the area (in square meter) between the two data also confirm that the two datasets overlap in space and that coupling the spatial and demographic information provided by the two different data-sets do not bring any systemic bias.

3.4.3 Results

Linear and super-linear scaling mirror the political history of the European Union

Europe Union is made of two major groups of countries, the Western European countries (WEu) the former members of the Eastern Block (EEu). The EEu countries were aligned to former USSR before its collapse in the late 1980s. Fig 3.22 shows the geographic partition of these two blocks. EEu countries experienced a transition from a totalitarian to a liberal economic regime. Theories and debates on the dynamics of the transition from totalitarian to liberal economy are wide [143]. The main expectation for transition economies is a fast economic growth given by the liberalization of the market, private and foreign investments and an emerging private bank system. Transition economies may eventually align to developed economies, like in the case of China with Japan. The WEu and EEu's economies perfectly fit the portrait of two convergent economies and they are a perfect case study to test universality of GMP scaling. Fig 3.21a shows the double peaked distribution of per-capita GMP for all cities. With few exceptions 3.21c its double peak shape overlaps the geographical East-West partition. The partition threshold is 11400 Euros per capita. Considering GMP vs population super-linear scaling independent of per-capita GMP, we wouldn't expect the double peaked distribution of cities' wealth effecting it. But EEu and WEu cities have very different scaling regimes with population. The exponent β of GMP on population for the EEu is > 1 and always major than in WEu cities, it increases over time from $\beta = 1.25 \pm 0.25$ in 2005 to $\beta = 1.42 \pm 0.20$ in 2010. On the other hand, in WEu cities $\beta = 1 \pm 0.03$ for the entire period, results are reported in Figure 3.21b. Linear scaling ($\beta = 1$) may mirror a stable and mature economic condition, as in EEu countries, while super-linear scaling ($\beta > 1$) represents a growing and

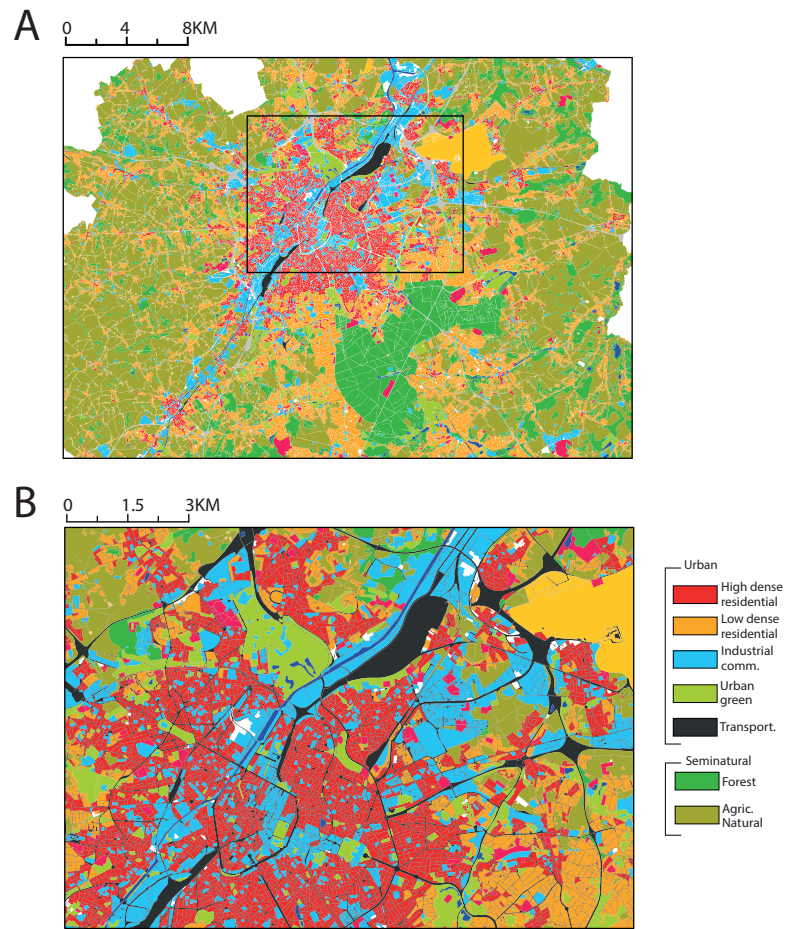


Figure 3.20: **Urban land-use example** A) A close-up on Bruxelles land use, colours indicate the different land uses partied by urban and semi-natural uses. In B) it is possible to observe that the class transportation do not only include streets but also the dedicated land.

3.4. Scaling of urban economic performance over time and its relationship with the urban land-use

unstable condition, typical of emergent economies such as the EEU one. A simple descriptive model describing the observed double scaling regime occurring in this process of economic convergence, is proposed. That the scaling exponent of GMP versus population increases, indicates that the GMP rate of growth is proportional to its population. Indeed if a city with a larger population will grow faster than a city with a smaller population, the GMP vs population curve (on a log scale) will become steeper over time and this is what we observe for EEU emerging economy. Given a city i with log-population p_i , let $g_{i,t}$ be its log-GMP for year t . We assume at this stage that the population of cities does not change to the same degree as its GMP. This assumption works well with observed data. For our analysis we used log scale for GMP (g) and population size (p). In log scale, scaling of GMP versus population size can be given a mathematical form:

$$g_{i,t} = \alpha_t + \beta_t p_i \quad (3.22)$$

where the exponent β tells us if the relationship is sub/super linear or linear. Empirically we have observed that β is 1 for WEU cities, and increasingly larger than 1 for the Eastern. The growth of the GMP can be modelled as a change from year to year,

$$g_{i,t+1} = g_{i,t} + \gamma_0 + \gamma_1 p_i. \quad (3.23)$$

Economic models of GMP growth have assumed a constant, population independent value of GMP growth, which would amount to $\gamma_1 = 0$. Plugging the GMP-population scaling relationship into the GMP-growth model, we find a relationship for the change in the scaling exponent,

$$\beta_{t+1} = \beta_t + \gamma_1. \quad (3.24)$$

A population independent GMP-growth (where $\gamma_1 = 0$) would cause the scaling exponent to remain stable over time, an observation we have made for Western cities. For eastern cities we observe a β_1 that grows for cities that are poorer than the West, and β_1 that is equal to 1 and stable for as rich as those in the West. To accomodate this last observation, we make the GMP growth rate depend on the GMP-itself resulting in,

$$g_{i,t+1} = g_{i,t} + \gamma_0 + \frac{\gamma_1 p_i}{1 + \exp(\theta(g_{i,t} - p_i - g_\star))} \quad (3.25)$$

Here the population size dependent term in the growth rate also depends on the city's GMP such that this term is smaller when the GMP per capita of the city is larger than a threshold value of γ_\star . This new form of GMP growth reduces to the earlier expression for cities that have a small GMP ($g_i \ll g_\star$).

When the city is rich, as is the case for all cities in the West, and a handful in the East, the population dependent growth term drops out. For large times, the cities keep growing until their GMP per capita crosses γ_* . At very long times, in a stable condition of economic convergence, we would expect all the cities to have crossed this threshold of richness, and the growth term be a constant and not dependent on the population. The attractor of the growth equation is,

$$g_{i,t} = g_* + p_i \quad (3.26)$$

implying that after a very long time we should expect all cities to show a linear scaling with population. This apparently simple figure might be prone to several speculations.

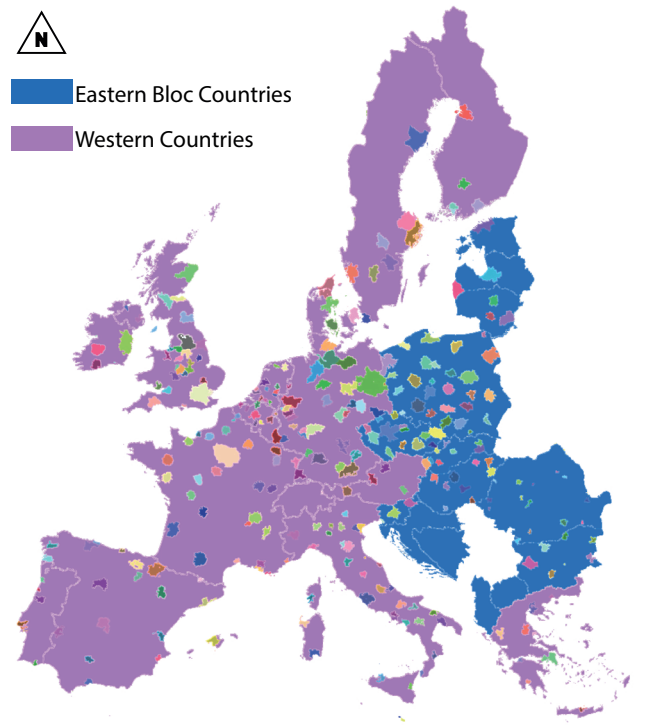


Figure 3.22: **West East partition.** The locations and the extents of all the cities case studies and the partition to former Eastern Bloc countries (blue) and Western European (purple) countries.

Land use scaling vs. double economic scaling regime.

Cities are a social and economic phenomenon but they are ultimately spatial events. It is worth to ask if and how the spatial organization of a metropolitan area is related to its economic performance.

3.4. Scaling of urban economic performance over time and its relationship with the urban land-use

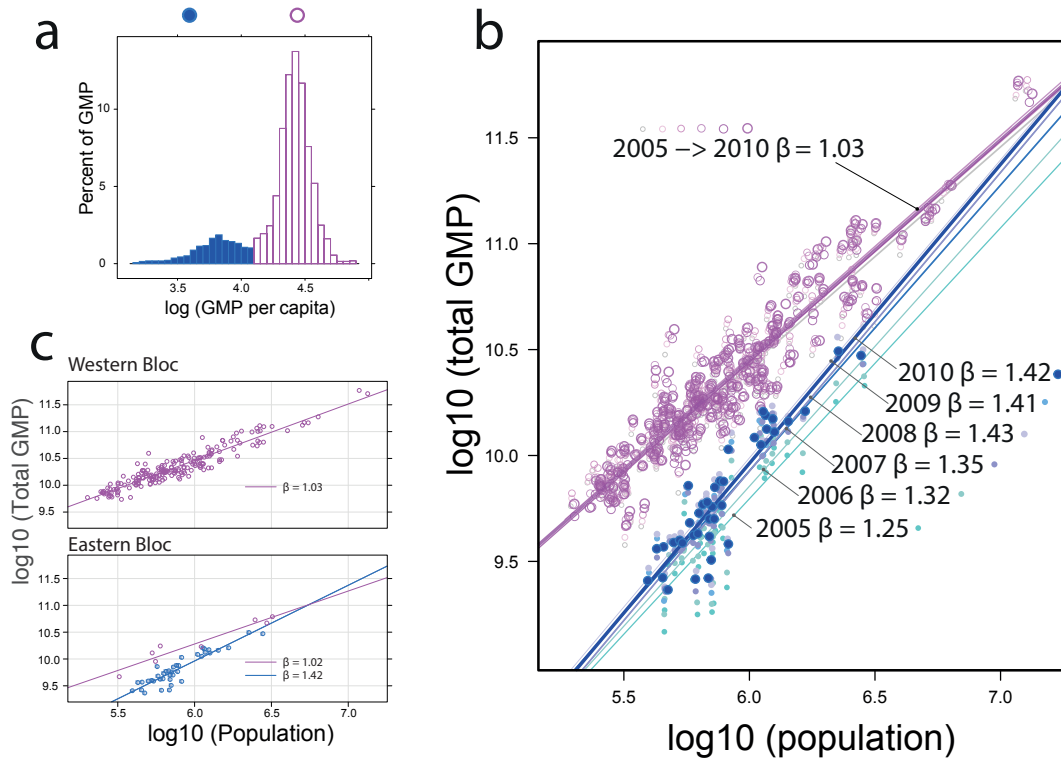


Figure 3.21: Linear vs super linear scaling. (a) Histogram of the gross metropolitan product of the 246 sub-samples cities. It is evident the double bell nature of the histogram indicating two classes of economies we called low income (blue) and high income (purple). (b) Time scaling of GMP vs population of both economies between 2005 and 2010. High income EEU cities show a stable linear exponent $\beta = 1.03 \pm 0.05$ while low income WEU cities show a super-linear with an increasing trend moving from 2005 $\beta = 1.25 \pm 0.25$ to 2010 $\beta = 1.42 \pm 0.20$ confidence interval for all years is 0.20 to 0.25. This shows that the two economies are differentiating more during time and that lower income cities in Europe are increasingly growing faster than high income cities. This makes sense also in the wider figure of global urbanization in which developing countries are growing faster than the developed ones. (c) Shows the scaling relationships of GMP and population for the year 2010 divided former Eastern Bloc (bottom panel) and former Western bloc (upper panel), the colours correspond to the map in Figure 3.22. It is possible to observe that no cities with low income belong to Westerns' plot, while only few cities with high income belong to former Eastern bloc cities. The overlapping low income/Eastern bloc cities high income/western bloc demonstrates that the geographical/political position of a group of cities overlaps with different scaling exponents and different trends and scaling.

Obviously any industrial city, with high regime of productivity, is expected to host more industrial land and to be more efficient in term of transportation networks. The dualistic behaviour of GMP in European cities offers a good sample to test this hypothesis. We know that spatial urban measures such as transportation surface or electrical cable are sub-linear with population, meaning that big cities require less road than smaller one. Sub-linear regime of spatial urban features is given by spatial constraint coupled with the planar nature of transportation networks. We measured industrial and transportation surfaces scaling with total urban area with population and with GMP. The natural question is if and how the duality found for GMP/Pop scaling is also mirrored in the purely urban spatial organization. Moreover, with the duality in the GMP values, is possible to detect biases in previous approaches that observe scaling of all cities in a single time without acknowledging different kinds of economies. In order to test this hypothesis three main classes of urban land use have been extracted from the Euro Atlas: 1) *Residential (Res)*: defined as the sum of the classes ‘continuous urban fabric’ (> 80%), ‘discontinuous dense urban fabric’ (50% – 80%), ‘discontinuous medium density urban fabric’ (30% - 50%). 2) *Industrial and commercial (Ind)*: defined as reported from original data. 3: *Transportation networks surface (Net)*: defined as the sum of the classes fast transit roads and associated land, other roads and associated land and railways and associated land. All this quantities have been associated to total urban areas (*TotUrb*) defined as the sum of all of them plus urban green areas and low-dense residential areas. Figure 3.20 shows the defined land-use classes for the city of Brussels. If compared to scaling in biology, *TotUrb* may represent the total mass of an organism while different land-uses classes may be seen as the mass of a given organ. So that studying the their scaling relationship is possible to detect the existence of a sort of ‘metabolic urban rate’. Scaling analysis have been performed using the entire set of cities at once and dividing West and East economies. In this way is possible to observe if and how different economies are associated with different urban metabolic rate. Results of such analysis are shown in Fig. 3.23. Scaling of the three land uses vs. *TotUrb* shows very stable exponents: for *Res* vs *TotUrb* $\beta = 1.02 \pm 0.07$, for *Ind* vs. *TotUrb* $\beta = 0.98 \pm 0.03$ while for *Net* scaling is, as expected, sub-linear and $\beta = 0.79 \pm 0.11$. In the case of Residential and Industrial, scaling for East and West Europe is stable around $\beta = 1$, variations around 1 are not sufficient to justify any super or sub-linear scaling. Transportation networks scaling, instead, presents different scaling exponents for the two economies. In East and developing economies $\beta = 1.06 \pm 0.07$, while for West and high income economies $\beta = 0.87 \pm 0.06$, as shown in the inset of Fig.3.23C. Linear scaling may represent an abundance and a saturation of transportation surface and it can be interpreted as a presence of redundancies in the transportation network. Linear scaling may imply in fact the existence of a regular and translation transportation surface, like for example a 2d regular grid. Sub-linear scaling, as for developed economies, reflects a possible dendritic nature of such networks in which larger the basin smaller the surface, and possibly a stage of optimization of the transportation network. Transportation network scaling also reflects the double GMP vs Pop scaling, suggesting that transportation network abundance can be a key factor in promoting economic development.

3.4.4 Discussion

Linearity for EEU cities may reflect the maturity of EU economy, as well as the strong European governance that redistribute incomes that aims to ameliorate disparity between small and large cities. Considering recent investment on internet connectivity and international high speed trains, and the

3.4. Scaling of urban economic performance over time and its relationship with the urban land-use

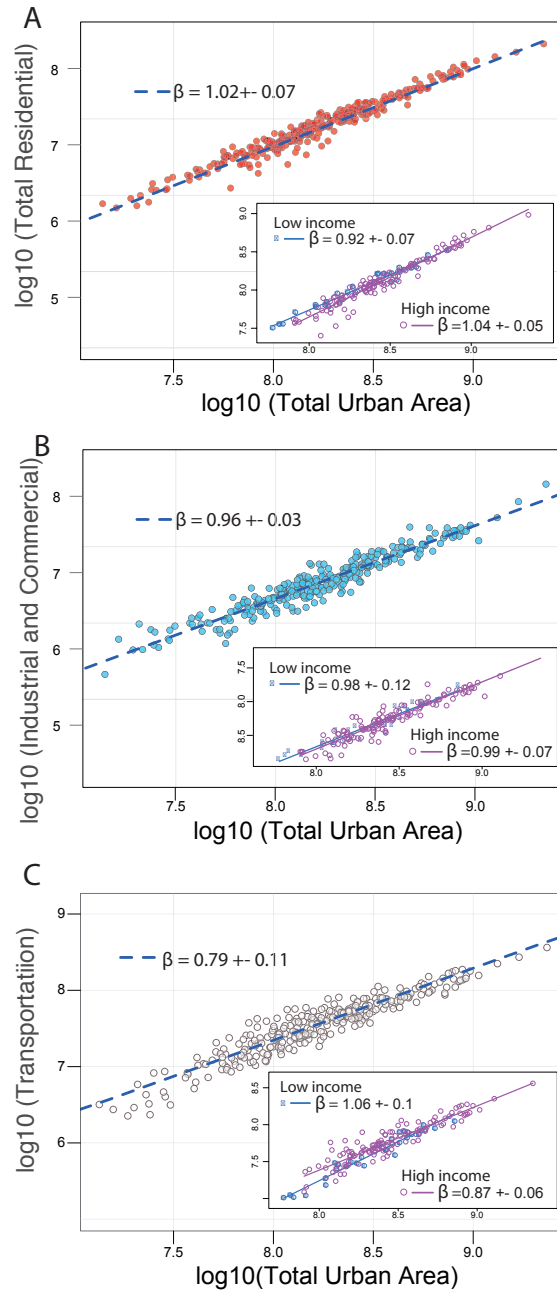


Figure 3.23: **Scaling relationships between main urban land uses and the total urban area.** Each dot in the plots represents a city. In all the plot the inset represents the same relationship as in the main plot but divided by low and high income groups. Plot in the inset, specially for low income cities have few points, so results must be taken carefully. (A, B) The scaling of total residential and industrial area are linear with total area and do not differentiate between low and high income cities. C) Scaling of total area of all transportation networks including local and fast roads and train, its shows a sub-linear behaviour, however it differentiate with economies' groups: it seems that for high income cities the exponent $\beta = 1.06 \pm 0.1$ suggesting an abundance of land dedicate to transportation networks in such cities, while instead for low economy β stays sub-linear.

increasing proximity of European capital cities, it may also be that for EEU economic activity an average individual can increasingly interact with out-of-town economic agents. As a result, the GMP of different cities can get too inextricably interconnected to satisfy the assumption of independent LUZ (or MSA). In the limit where an individual can interact with anyone in the larger economy of a state, her economic output will depend on the total population of the larger economy, and thus be independent of the population of the city where she resides. This will result in the traditional linear GDP vs population size relationship. Another observation supporting linearity of GDP with population as a signature of a mature economy is the exponent for the richer Eastern cities. The exponent for these cities is closer to 1, than for the Eastern bloc as a whole (Fig.3.21). Thus the exact scaling of GDP with population does not depend only on the geographic/historical context, but also (may be solely) on how rich the city is.

The situation for Eastern and emerging LUZ tell us different story. After USSR collapse these economies have been developing fast, and more populous cities have been developing faster and faster over time. Expectedly more populous cities have seen higher rates of growth causing the exponent β to increase from 1.25 to 1.42. Increasing β is a signature of increasing economic performance but it also means increasing inequalities between cities. In a way, the observed super-linear scaling might reflect an unbalanced situation of rapid growth of large town and economic segregation of small ones, situation in which the redistribution of income is increasingly harder. These findings are in line with recent considerations on the economic stuck of Eastern Europe and their increasing social inequality [144]. The increasing super-linearity can also represent an increasing delay in the economic convergence process between the two economies. Regarding scaling of economic performance of cities, it has been showed that it is a powerful tool to observe economic fluctuations over short time period. However its universality still remains difficult to prove. This in turn suggests that the increasing connections between cities, and not only inside cities, cannot be neglected for understanding urban dynamics.

Scaling between different land-uses reveals a strong regularity, suggesting a single and universal mechanism regulating urban growth. In this sense, all cities seems to be just a scaled copy of a typical city. Regularity and patterns in the spatial organization of towns, however, is a contrast with the dualistic economic behaviour found for GMP/POP scaling. This findings suggests that spatial organization of towns does not effect their economy, at least for the considered cases. Scaling of transportation network instead, is somehow mirroring the dualistic GMP/Pop scaling. Low-income fast-growing economies seems to have an abundance of transportation surface if compared to high and stable-economy ones, in which transportation surface may have found an optimal form.

3.5 Structural proprieties and scaling of the global road network

The study proposes the analysis of the network composed by all roads on Earth. The study is a part of research output of the research project *Exploration in Urban Scaling* funded by ENAC, EPFL. The results presented in this section will be submitted as self-contained scientific publication. The paper under preparation is:

Strano E, Bertuzzo E, Shay S, Giometto A, Mucha P, Rinaldo A, *Structural proprieties and scaling of the global roads network*. in preparation.

Contribution of the candidate to the study include: research design, preparation and analyses of the data, analysis of the results and writing of the manuscript.

3.5.1 Introduction

All modern civilizations developed along with an efficient road network. A road network is a simple and highly efficient system for colonizing free land, improving communication and moving goods among locations. Today, the road system fragments the 19 million hectares of the Earth's surface with more than 14 million kilometres of paved surfaces. As the backbone of the human colonization of Earth, road expansion embodies the complicated cohesion of economic growth and sustainable development [145, 146]. From an economic perspective, road expansion has typically been associated with economic growth, reduced poverty [147] and urbanization [148]. However, roads cause severe environmental degradation: habitat disruption, urban sprawl, and deforestation are a few of the side effects of road expansion [149]. In the upcoming decades, considering the global population growth [2] along with the resulting increase in food demand [10] and global urbanization, massive road expansion will be inevitable. It has been estimated that the total paved surface length will increase to 25 million kilometres by 2050 [150]. Controlling such an expansion is of crucial importance for global environmental conservation strategies and sustainable agricultural development [151]. However, despite the fundamental role of roads in the global human-environment relationship and some attempts to reconcile the dual effects of road expansion [151], a quantitative and objective description of the structure of road networks is still lacking. Here, coupling a detailed and unprecedented dataset on the global road network (GRN) [92] with global land-use inventories (see Chapters 2.1 and 2.2), an analysis of the structure of the major road networks on Earth in 2015 is provided. The analysis sheds light on the universal laws that govern the process of global land fragmentation. Such laws are defined using road length probability distribution and scaling analysis. The analysis was performed observing two main characteristics of the GRN: 1) the land use each road belongs to and 2) the hierarchical road structure.

As reported in Chapters 1.3 and 3.1, physical laws governing road networks have been extensively explored in the past [152, 70]. Such studies investigated problems such as scaling [65], centralities [73] and evolution [108, 120] of roads. However, such studies exclusively focused on urban road networks, never examining the global figure. Moreover, correlations between the structure of road networks and extra-urban land uses remain unexplored. This case study fills this gap by providing evidence of a universal scale-invariant road structure that is independent of land use and location. In other words, the probability density functions of road lengths in urban, wild and agricultural land, if properly scaled, show the same functional form. It is shown that the self-similarity of land use classes could be caused

by the self-similar hierarchical organization of roads. The results suggest that the processes that govern road expansion are universal and that local conditions, such as economic development or land use, can accelerate or decelerate infrastructure development; however, over the long term, these processes cannot regulate the roads' overall length distribution or scaling. These results are relevant for two reasons: 1) they offer the very first structural analysis of the global road network, which provides a new view for studying and quantifying the human presence on Earth; and 2) evidence of universal scaling can play a crucial role in modelling future road development, which have been estimated to double in the next few decades. In this regard, universal scaling laws that regulate the recursive fragmentation of land can be of great importance in more reliable development strategies.

3.5.2 Materials and methods

The study uses two global urban inventories coupled with global land use inventories. The data were reported in Chapters 2.1 and 2.1; the materials and methods are reported below. The data preparation process includes the extrapolation of urban masks from Nightsat data and a series of spatial join operations; Figure 3.24 shows the full data process schema.

Global road networks.

In this study, the Global Road Network, as described in Chapter 2.1, is used. After cleaning operations and topological corrections, we defined the GRN as a *primal road network* [66], in which nodes (N) are the street junctions and links (E) are street segments. Each link carries a weight W that is equal to its length (l). The GRN does not contain urban local roads; it only contains major roads and is organized into four hierarchies: primary roads with limited access (H1), primary roads with non-limited access (H2), minor and secondary roads (H3), and local roads (H4).

Nightsat and urban mask.

Nightsat [39, 153] is continuous layer of 2.7x2.7 km resolution with values ranging from 0 to 63. Each value represents the intensity of stable night light in 2013. These data are publicly available and have been produced by NOAA's National Geophysical Data Center; DMSP data were collected by the US Air Force Weather Agency. To extract an urban settlement mask, the Jenks algorithm was applied [154]. The Jenks algorithm is widely used in geographical analysis to cluster continuous surface data sets. Jenks involves unsupervised classification methods that impose the number of clusters. From a set of randomly selected values, this algorithm minimizes the variance value inside classes while maximizing variance between classes. Several classifications with different numbers of clusters have been tested. Clearly, if the number of cluster is small (two or three), the classification will take only the very highly illuminated places and exclude medium illumination places. The higher the number of clusters is, the smoother the provided classification will be. However, considering the sharp separation of high and low illumination places, different classifications did not affect the final classification of urban masks and thus did not affect the analyses. Figure 3.24A and D shows this classification.

Global cropland inventory.

The 1 km global IIASA-IFPRI [95] cropland percentage map for the baseline year 2005 was developed by integrating a number of individual cropland maps at global to regional to national scales. The individual map products include existing global land cover maps such as GlobCover 2005 and MODIS v.5, regional maps such as AFRICOVER and national maps from mapping agencies and other organizations. IIASA-IFPRI is publicly available data, and it can be downloaded from the Geo-wiky platform at [96].

Data fusing and geographical partition.

Coupling the GRN with the night light urban footprint [39] and with cropland global inventory [95], each road has been sorted into three classes: roads fully belonging to urban areas (Ur), roads bounding or fully belonging to cropland (CLr), and wild roads (Wr), which are roads free of direct human presence, such as roads crossing natural parks or remote areas. Each road segment was also associated with the political boundaries of all the countries and continental macro regions. The macro regions are based on the geographical macro regions provided by the United Nation and are the following: North America, Central America, Caribbean, South America, Northern Europe, Western Europe, Southern Europe, Eastern Europe, Northern Africa, Western Africa, Middle Africa, Eastern Africa, Southern Africa, Central Asia, Western Asia, Southern Asia, Eastern Asia, Southeastern Asia, Melanesia, Australia and New Zealand. Further partitioning was performed using country boundaries. In the case of China, US, Brazil, Russia and India, state boundaries were used.

3.5.3 Results

The GRN in 2015 expands on Earth with a total length of 14,522,470 km with more than 3 million road segments. Considering only major roads ($H1 \cup H2 \cup H3$), the total length decreases to 10,095,163 km. This simple quantity is in sharp contrast to the gROADS database (see Chapter 2.1) recently used for global road environmental impact estimation [151] that gives a total road length of 7,644,410 km. Considering that gROADS reports the global road network in 2009 and assuming that the two data comparable, the annual rate of growth of major roads in the period of 2009-2014 is up to 5.7%. However, considering all road hierarchies ($H1 \cup H2 \cup H3 \cup H4$), previous analyses underestimated global road length by up to 30% and are thus likely to underestimate their environmental impact. The entangled cross-accuracy estimation of different spatial datasets precludes more accurate estimations of road evolution or mutual biases, but clear discrepancies between them warrant further investigations.

Ur, Wr and CLr cover 12%, 37% and 51% of the global road length, respectively. Roads shorter than 100 meters contribute only 0.03% of the total length (Figure 3.25B). Such roads, as confirmed by an extensive inspection, are highway ramps or road segments for large road junctions and are therefore excluded in the further analyses.

A probability density function of road length, as shown in Figure 3.25C, shows that CLr and Wr, despite their differing total length and their differing nature, almost overlap for the same distribution. This indicates very similar patterns of road lengths. In other words, roads fragmenting wild lands, such as natural parks, are very similar (in term of size) to roads built for agricultural or livestock purposes.

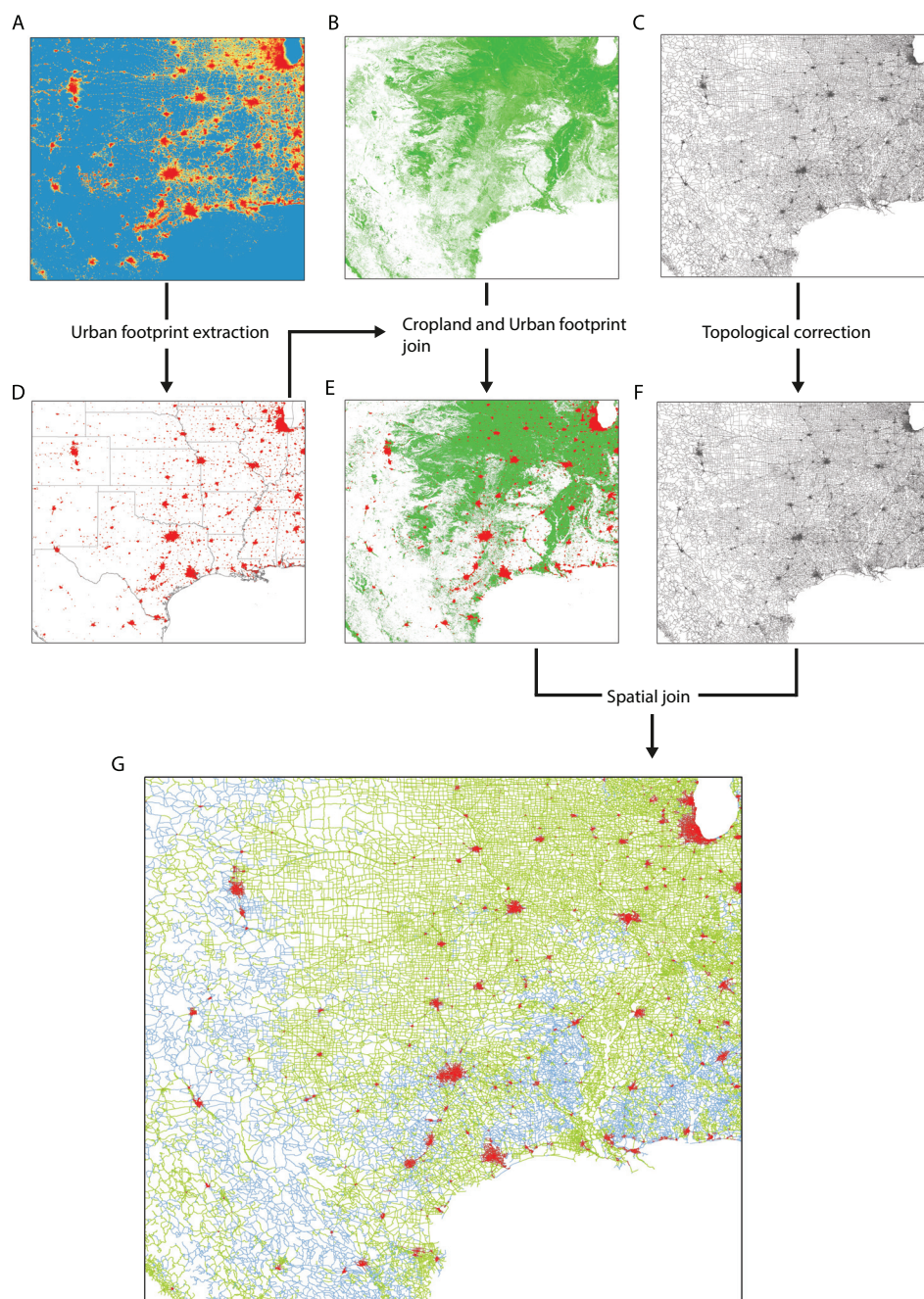


Figure 3.24: **Global road network data preparation.** (A) Nightsat data, stable light average. (B) Global cropland layer. (C) Global road network. (D) Urban footprint mask extracted using the Jenks clustering algorithm. (E) Cropland and urban footprint mask. (F) Corrected global road network; in this case, the corrections are not visible. (G) Final network divided into three classes.

Similar length distribution suggests a high level of fragmentation of areas supposed to be wild and free of human presence. Although there is a similar distribution in Wr , there is an abundance of longer segments, while the longest road belongs to CLr . This might appear counter-intuitive, but closer observation reveals that very long single road segments, for example, roads that fragment the Amazon forest, are surrounded by agricultural fields and therefore belong to the cropland class. This is a simple proof that the land fragmentation process as a result of wild land fragmentation is because of wild to agricultural land conversion. As expected, Ur segments (Fig.B) are shorter than those of CLr or Wr ; however, the functional form is similar to that of Wr and CLr . The same functional form for Ur , Wr and CLr would imply a scale-invariant length distribution. To test this hypothesis, we tested whether $P(l)$ shows a finite-size scaling form [155, 156, 157]. Finite-size is obtained using an algebraic power of the length multiplied by a suitable scaling function F :

$$P_c(l) = \frac{1}{l^\gamma} F\left(\frac{l}{\langle l \rangle_c^\alpha}\right) \quad (3.27)$$

where $\langle l \rangle$ is the mean of the road length in the land use class c and the function $F(x)$ is the same scaling function for all land use classes. The finite-size form shown in Eq.3.27 implies that length distribution is dependent only on the mean length $\langle l \rangle$ and is thus independent of the land use. As demonstrated in [158], the two parameters α and γ are not independent, and they must satisfy the relationship $\alpha = 1/(2 - \gamma)$. The best collapse is traditionally found by visual inspection of probability density plots for different α . Here, we employ a quantitative method, as proposed in [159]. Tuning α in a suitable range between 0.5 and 1.5, the quantity $E(\alpha)$ measures the cumulative area, within the same support, between curves. The value of α that best minimizes $E(\alpha)$ is taken as a collapsing exponent and is shown in Figure 3.25C (Inset). Plotting $l * l^\gamma$ vs. $l / \langle l \rangle_c^\alpha$ for the three classes and tuning the two parameters α and γ , a satisfactory collapse was found for $\alpha = 1.1$ and therefore $\gamma = 1.11$, as shown in Figure 3.25C. The same analysis repeated for the macro regions presents similar results: the range of best collapses for all regions is $1 < \alpha < 1.1$, as shown in Figure 3.26. The collapse of road lengths for different land uses is independent of location and can be considered a universal feature. Given the very different nature of land uses, the expected differences between locations, and the use of all roads on Earth, the observed scaling is compelling.

Because the road length distribution does not address the topological structure of GRN, we measure the quantity $e = E/N$ in all macro regions and for all countries. For a tree-like planar network, $e = 1$; for a regular 2d grid, $e = 2$. The results in Figure 3.27 show a universal value of $1.38 < e < 1.51$ among locations. These results indicate that all road networks on Earth have a very similar structure. Fluctuations of e on the order of 0.1 are too small to justify the diversity of road patterns among locations.

The results of the finite-size and topological analyses display road patterns that are similar across scale, location and land use. A possible explanation of such regularity could be rooted in the process of land use conversion. Is reasonable to imagine that urban land was previously dedicated to agriculture. Thus, the resulting lengths of major roads do not change over time. Another possible explanation could be the process of evolution and hierarchical fragmentation of roads. It is possible to investigate the fragmentation process by studying the relationship between the size of a road block of a given

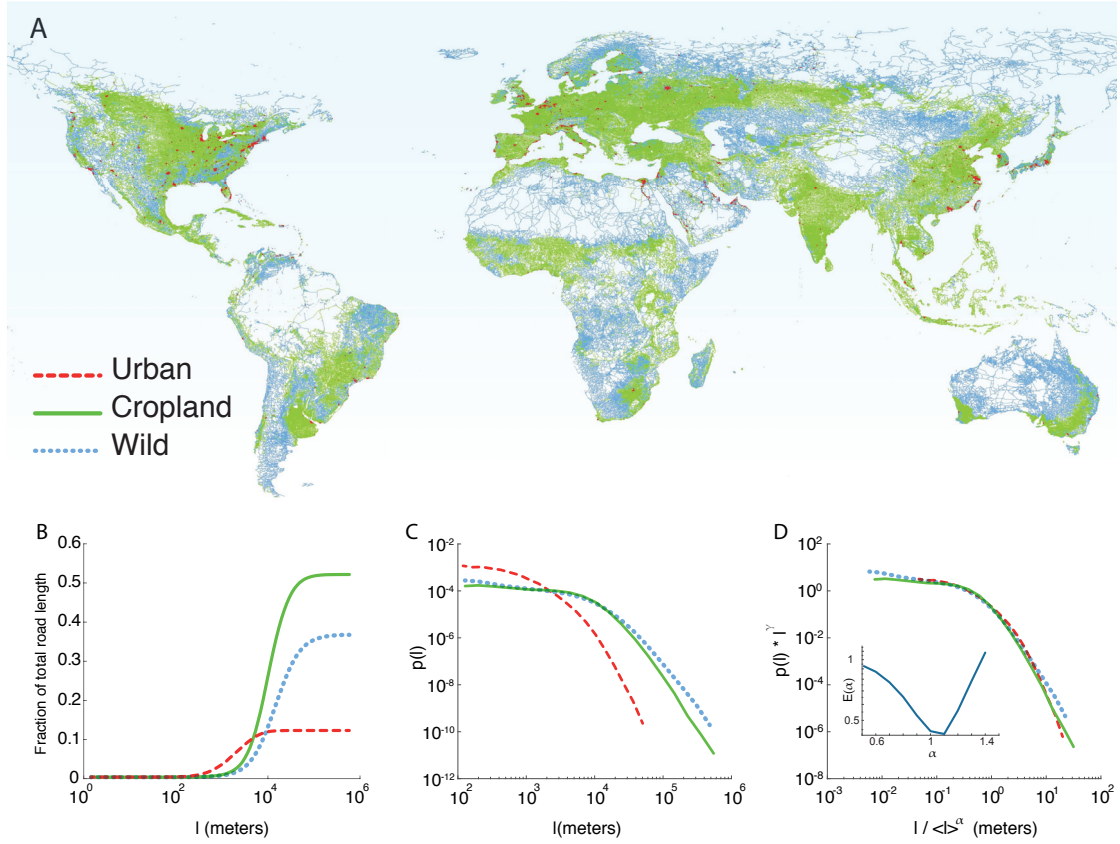


Figure 3.25: **Evidence for a universal length distribution.** (A) A visualization of the classified GRN; different colours represent different land uses: red is Ur, green is CLr, and turquoise is Wr. (B) Fraction of total road length composed of roads with lengths of less than l (x -axis) for each land use class. (C) Probability density function $p(l)$ of road lengths for each land use. (D) Collapse of length distribution obtained by showing $p(l) \cdot l^\gamma$ vs. $l / \langle l \rangle^\alpha$ (Methods), where $\langle l \rangle$ is the mean length. The best collapse was found for $\alpha = 1.1$ and $\gamma = 1.11$ (inset).

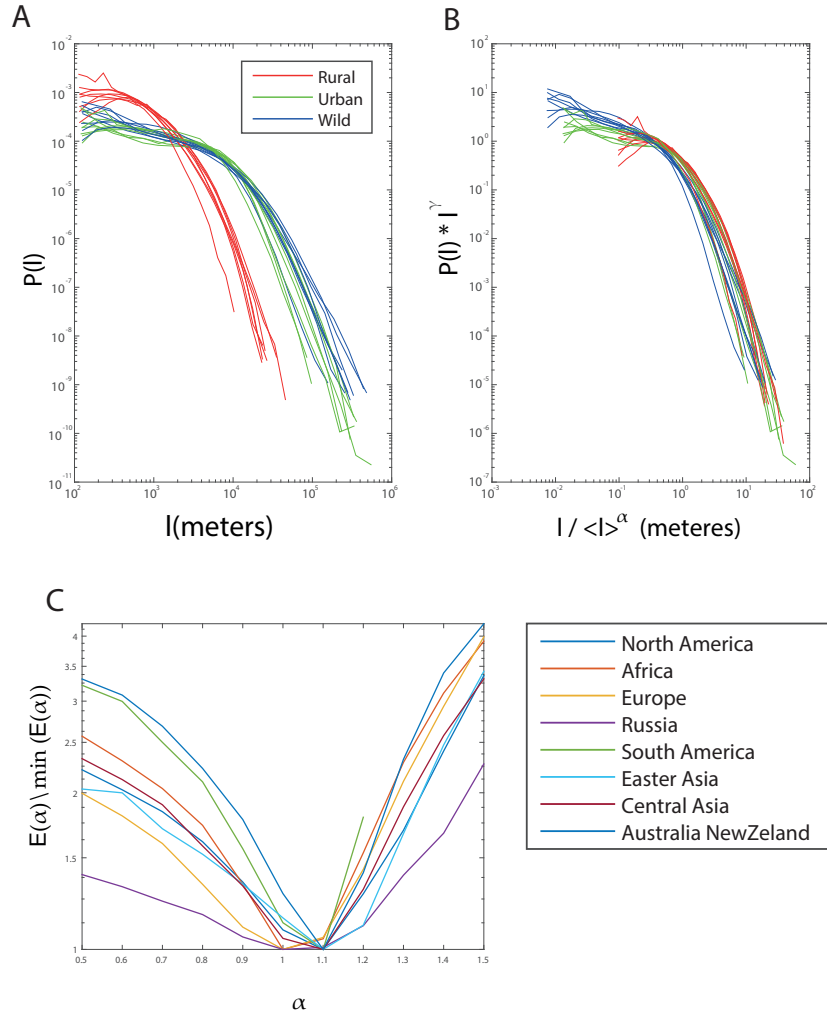


Figure 3.26: **Length distribution and scaling by macro regions.** (A) Road length probability distribution function for each land use and each macro region. (B) Collapse of length distribution obtained by $p(l) * l^\gamma$ vs. $l / \langle l \rangle^\alpha$. (C) The best collapse, which was found for $1 < \alpha < 1.1$ and $\gamma = 1/(2 - \alpha)$.

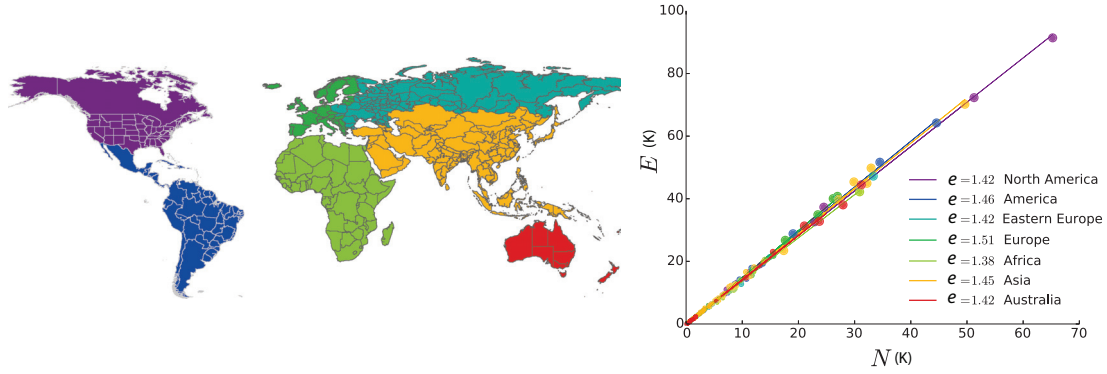


Figure 3.27: **Topological road structure for all countries.** The plot shows the quantity $e = E/N$ for all macro regions, as illustrated in the map. Each point in the plot represents a state (for US, Brazil, Russia, China and India) or a country. It is possible to observe a universal value for e of between 1.38 and 1.51.

hierarchy (or network face) and the length of roads belonging to it and to the lower hierarchy. For example, considering only highways (H1), it is possible to extract a series of large network faces and to then study how these faces are fragmented by major roads (H2). Then, in the same manner, we can observe how the faces of major roads are fragmented by local roads (H3) and so on. Figure 3.28A and B illustrates this idea. If any recursive and scale-free fragmentation process exists, regularities in the road length distributions for all hierarchy partitions should arise and must be independent of the area of the network face. This hypothesis has been tested again using finite-size scaling. After dividing all road segments by their hierarchical class $E_{H(i)}$, each of them was assigned the area $A_{H(i-1)}$ of the network face they belong to. Then, it is possible to define a vector of road length, such as $l_{A(k)} = \{l_{H(2)}, l_{H(3)}, l_{H(4)}\} \cap A_{(k)}$ that contains all road lengths that belong to the area range k . To ensure a cross-hierarchical analysis, l_A includes all hierarchies. We can then study the relative proportion $p_k(l)$ that measures the probability of finding a road of a given length in the area range k . We observed again that if the scaled probability $p_k(l) * l_k^\gamma$ vs. $l/\langle l \rangle_k^\alpha$ exhibits a finite-size scaling behaviour, this would imply a self-similar and cross-hierarchical fragmentation process. Using the full dataset, a satisfactory collapse was found for $\alpha = 1.1$ and for $\gamma = 1.1$, as shown in Figure 3.28. Such a collapse indicates a universal scaling curve that regulates the fragmentation of road network faces. $\alpha = 1.1$ indicates that larger faces are more fragmented than smaller ones. The universal curve depends only on the size of the network face and not on the hierarchy itself. Fragmentation pressure on network faces is universal and operates globally.

3.5.4 Discussion

This study presents the very first analysis of all major road networks on Earth. The GRN is a fully connected and huge infrastructure that represents, together with cities, the human footprints on Earth. GRN spans over the entire globe, leaving very few empty zones, which are basically mountain and desert areas. The analysis focused on the physical structural characteristics of the GRN, which display self-similarity among land uses, locations and hierarchical structures. Self-similarities have been tested using finite-size analysis on road length segments. We demonstrated that road lengths in cities, cropland and wild areas are distributed following a universal scaling function. The universality of the

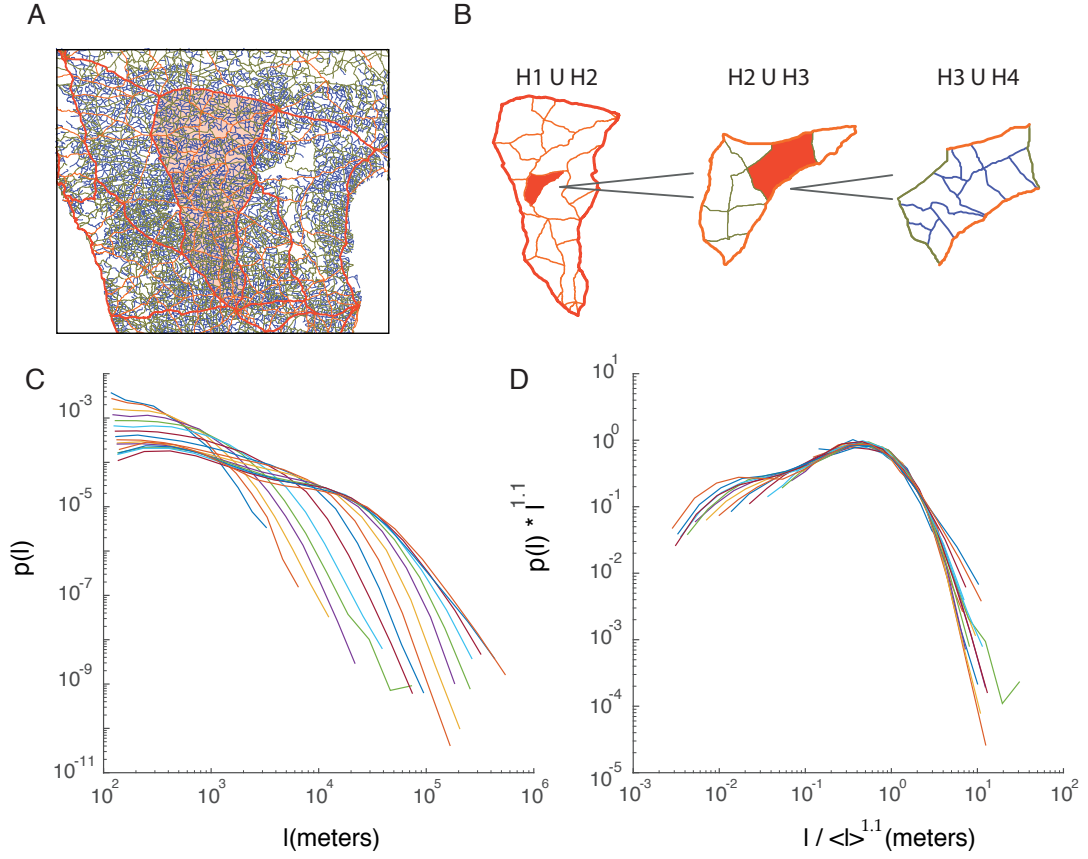


Figure 3.28: **Hierarchical scale-free fragmentation of the global road network.** (A) The Indian road network in which the colours from red to green represent H1 to H4. (B) The process of recursive hierarchical fragmentation in which, starting from H1, each face is fragmented by the link of the lower hierarchy. (C) Probability distributions $p(l)$ of the link belonging to a block area range k . (D) Collapse of the road length; the best collapse is found for $p(l) * l^{1.1}$ vs. $l / \langle l \rangle^{1.1}$.

scaling function implies that the rules that govern the GRN expansion are independent of the land use. The universality of the scaling function can be related to the results presented in Chapter 3.1, in which urban and peri-urban road networks are governed by simple spatial roles. Such regularity can be also rooted in the fragmentation process of road network faces that shows scale-free behaviour. The results of the physical structure of the GRN have important implications for understanding the physical laws governing global human footprints. In particular, it has been proven that the global process of land fragmentation, which is imposed by the construction of roads, can be seen as a gigantic unique process of fragmentation that operates beyond local constraints. Evidence of a unique and single operating global fragmentation process, which displays self-organized characteristics, poses relevant questions on the role of local policies that regulate human colonization of Earth and, in particular, road expansion. The effectiveness of any road expansion regulation depends on its ability to control the entire process across multiple scales. However, such an approach has yet to be implemented by local authorities, which in most cases, impose only local rules. Awareness of cross-scale fragmentation pressures can be also fundamental to understanding the implications of land fragmentation of wild areas. For example, as already suggested by a few studies on the ecological impact of new roads [160, 161], even a single road crossing a wild area can activate a process of further land fragmentation, thus posing potential damage to ecosystem conservation. Here, we provide a quantitative analysis of qualitative observations. The pressure of fragmentation of network faces are independent of hierarchy and depends solely on the network face areas. Other observations are related to the impact of a self-similar transportation structure on the process of global urbanization. Scale-free urbanization patterns, as described in Chapter 2.1, can be due to the road network structure and not because of any intrinsic urban quality. This speculation is corroborated by the universal and land use-independent functional form of the road length distribution. To the extent to which a road pattern is preparation for further urbanization, most future urbanization is already there.

Finally, this study provides a new vision of global human footprints in which the “‘network space’ is taken into consideration along with the land use. Such a vision needs further developments, however. Considering the global urbanization studies in perspective, the present study represents a foundation to build upon.

3.6 Testing the universality of urban scaling.

The study proposes the analysis of scaling law for all urban settlements on Earth. The study is an output of the research project *Exploration in Urban Scaling* funded by ENAC, EPFL and in participation with DLR (German Aerospace Agency).

The results presented in this section are mostly preliminary and preparatory results. Final results will be submitted as three self-contained scientific publications. The papers in preparation are:

1: Strano E, Rinaldo A, *Typologies of urbanization patterns, a global sample*. In preparation

Candidate's contribution to the study includes: design of the research, preparation and data analysis, writing of the manuscript.

2: Esch T, Strano E, Heldens W, Hirner A, Keil M, Marconcini M, Roth A, Zeidler J, Rinaldo A, *The Global Urban Footprint Initiative - Pushing the Limits of Mapping Global Human Settlements from Space*. in preparation.

The candidate participated to the production of the global data set, in particular the candidate worked to set a global data validation framework and wrote the paper.

3: Esch T, Heldens W, Hirner A, Keil M, Marconcini M, Roth A, Zeidler J, Rinaldo A, Strano E *The global urbanization profile - A worldwide inventory of human settlements*. in preparation.

The candidate conceived the production of a global and public available human footprint inventory, prepared the data and wrote the paper.

3.6.1 Introduction

The pervasive scaling nature of cities and urban patterns has been extensively documented in the past, as described in Chapter 1.2. As stated previously, the scaling and self-similarity of cities are mainly of two natures: spatial, such as the classical scaling of urban built-up masses, and non spatial, such as the scaling of gross domestic product vs. population. In this section, using a global sample of urban settlements, the hypothesis of spatial urban scaling has been explored. The main aim is to validate or reject, with an empirical analysis, if and how the scaling law for urban settlements is a universal and cross-cultural feature. The specific hypothesis then pertains to the universality of the scaling law (better known as Zipf's law) is expressed as $P(x) = Cx^{-\alpha}$, where x is the area of an objectively defined urban mass and $\alpha = 1$. To strictly ensure spatial analysis of urban patterns, urban masses are defined as contiguous built-up areas (or urban patches) and not as political entities. Political boundaries of towns are subject to political definitions of cities, which are diverse among regions and cities. Defining urban masses as pure spatial entities avoids the biases from regional subjectivities in city definitions.

This study was motivated by three specific knowledge gaps: 1) Most of the studies on urban scaling have been focused on single metropolitan areas centred on a large city. Measuring scaling laws in mono-centric urban areas may have several side effects caused by the limited range of sizes. In practice, bounding the study area around single cities reflects the urban form but reflects very few patterns of

urbanization at higher scales that are, in most cases, composed of thousands of medium-small urban sites. 2) The scaling law approach led to homogenization of the urban phenomena in which all urban locations have been associated with a single scaling parameter, typically $\alpha = -1$. This is a fundamental law that may not be exhaustive of the potential complexity of the urbanization phenomena. More insights are necessary to track the boundaries between universality and diversity of urbanization patterns. 3) Lack of high resolution and updated global urban inventories precludes a precise scaling analysis. Because scaling laws are subject to a lower cut-off, below which X rapidly goes to zero, resolution of remotely sensed data (that regulates the minimum urban site size) could play a crucial role in objectively defining the lower cut-off and the real extent of the power law distribution.

In this study, the scaling analysis uses two remotely sensed spatial urban inventories: the GL30, at a quasi-global level, and the GUF for a subset of the four large regions of NY state, the Sao Paulo metropolitan area, Niger and South Africa. Both datasets are described in Chapter 2. The results of this section are organized as follows: First, scaling of a global dataset, which is derived from GL30 and includes almost all settlements on Earth, is analysed. Second, scaling of urban settlements was tested on a regional scale and for all countries. Regional scaling shows a non-universal behaviour in which many urbanized locations do not present power law distributions. Third, a comparison between low and high resolution urban spatial inventory is presented. This comparison shows that the scaling approach could be used to differentiate urban and rural patterns and that the resolution of the data is a critical feature in a scaling analysis. Fourth, the new high resolution global urban inventory derived from GUF is introduced together with concluding remarks and perspectives.

3.6.2 Results

Considering most urbanized areas on Earth (Figure 3.29A), 1,117,588 urban patches have been extracted and measured from the GL3 data. The areas of these patches span 6 orders of magnitude from a minimum size of 3600 m^2 to a maximum of 4652 km^2 (Figure 3.29B). The largest urban patch on Earth is the urban agglomeration that includes Los Angeles and San Bernardino in California. The second largest urban patch is Chicago, Illinois. The cumulative distribution plot of the size of all urban settlements, as shown in Figure 3.29B, shows a clear power law distribution with the exponent $\alpha = -1.17 \pm 0.02$. A coarse-grained analysis, which tests the stability of α for lower resolution sub-samples, confirms the self-similar form of the cumulative distribution, as shown in Figure 3.29C. This simple plot confirms that all settlements on Earth are distributed as a power law with an exponent close to 1, thus confirming that Zipf's law works at a global scale. This figure is open to several speculations.

The scaling of urban mass' areas at the regional scale, however, presents a different figure. For the global road network analysis (Chapter 3.5), regions have been defined by countries' political boundaries; for the US, Brazil, Russia, China and India, state boundaries have been used, as shown in Figure 3.29A. Clearly, any long tail distribution can be fitted by a power law, but the analysis of *Pvalue* [63], which quantifies the plausibility that the urban masses follow a power law distribution, shows that in many regions, the probability distributions of urban masses are not compatible with a power law. Most notable examples of power law incompatibility are highly urbanized countries such as UK and France and heavily populated Chinese regions such as Shandong and Shaanxi. The explanation of this incompatibility could be rooted in the presence of a few large metropolitan areas together with "typical"

cities of medium or small size. Another possible explanation, especially in the case of China, could be the hyper-urbanization. Areas such Shandong and Shaanxi have more than 50% urbanized land. More analyses are needed to verify these preliminary tests. In particular, different functional forms could be tested in comparison with the percentage of urbanized land.

A comparative test between high and medium resolution data shows that the scaling approach is able to discriminate between urban and rural settlements. In four large areas, namely, Sao Paulo, New York state, South Africa and Niger/Nigeria, the scaling of urban settlements was done using both high resolution GUF data (see Chapter 2.1) and medium resolution data (for previous analyses). The results of this analysis are presented in Figure 3.30 and in Table 3.2. The main hypothesis pertains to determining how the data resolution affects the lower cut-off of the power law distribution. The lower cut-off can then be related to the presence of different distributions and, in particular, an abundance of small and rural settlements. For all case studies, the lower cut-off was found using the method proposed by Clauset [63]. For the New York, Sao Paulo and South Africa case studies, the lower cut-off $\sqrt{X_{min}}$ decreased for GUF data closer to the area of the smaller settlement $\sqrt{min(A)}(m)$. In practice, improving the resolution confirms that the power law distribution spans over the entire size range, from the smaller to the larger urban patches. In the case of Niger, $\sqrt{X_{min}}$ increases for the GUF data. Niger is the only case in which more detailed data create a larger lower cut-off. A map of settlements above and below the lower cut-off, as shown in Figure 3.31, shows that settlements below the threshold are mostly rural settlements. It is very interesting to note that settlements below the lower cut-off are not connected by any paved transportation network. This suggests that road networks, which increase connectivity between settlements, could play a crucial role in determining a hierarchical and self-similar urban structure.

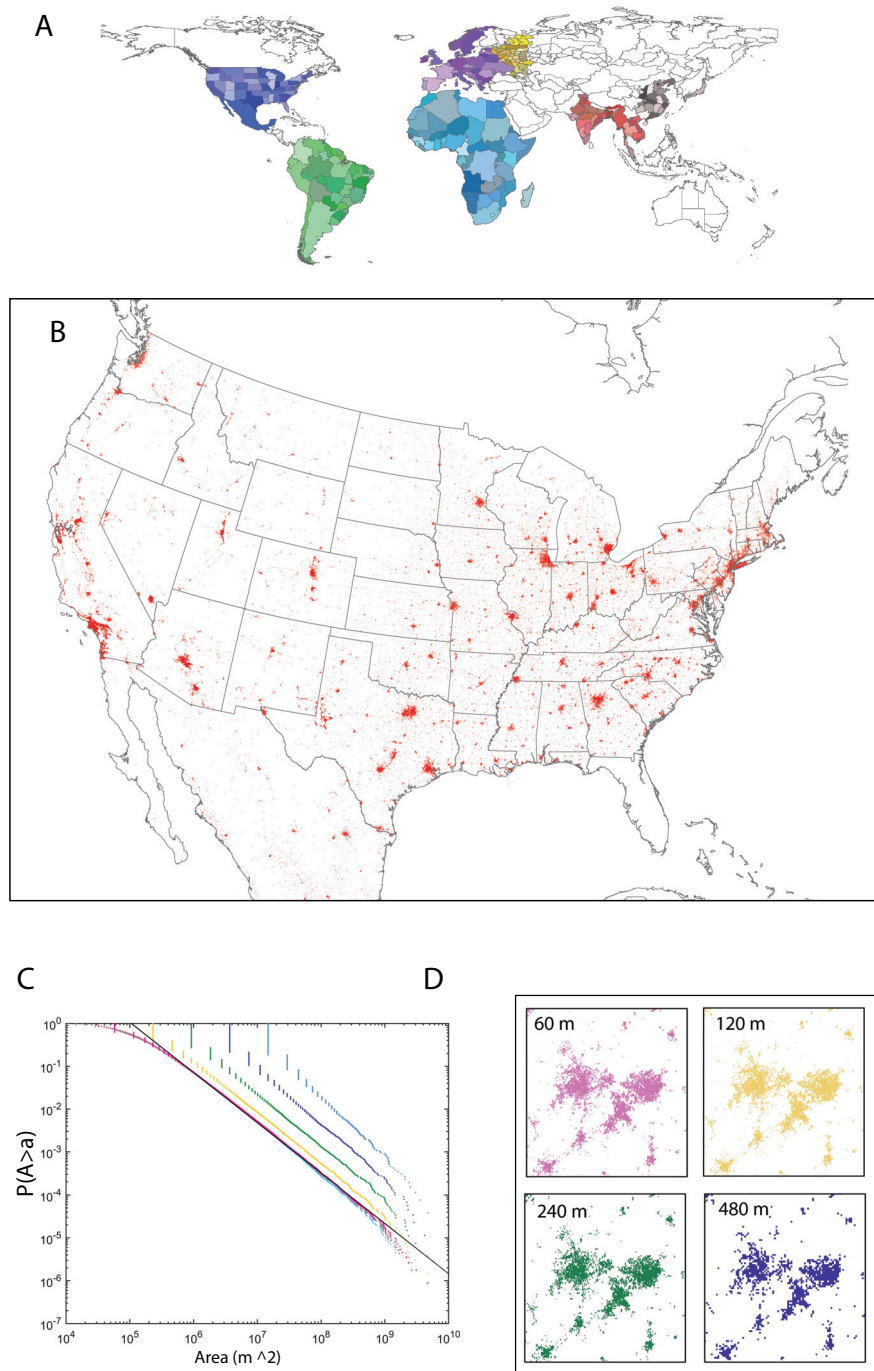


Figure 3.29: **Scaling of all settlements on Earth.** A) The maps show the extent of the analysed areas; different colours represent different macro regions. For the US, China, Brazil, Russia and India, the regions were divided using state boundaries. B) A visualization of the US urban settlements. C) Cumulative distribution plot of the areas of all settlements. Different colours represent the full data at different resolutions. It is clear that the lower the resolution is, small settlements disappear; however, the slope of the distribution remains $\alpha = -1.17 \pm 0.02$. D) The maps correspond to different resolution tests; the colours correspond to the plots in C.

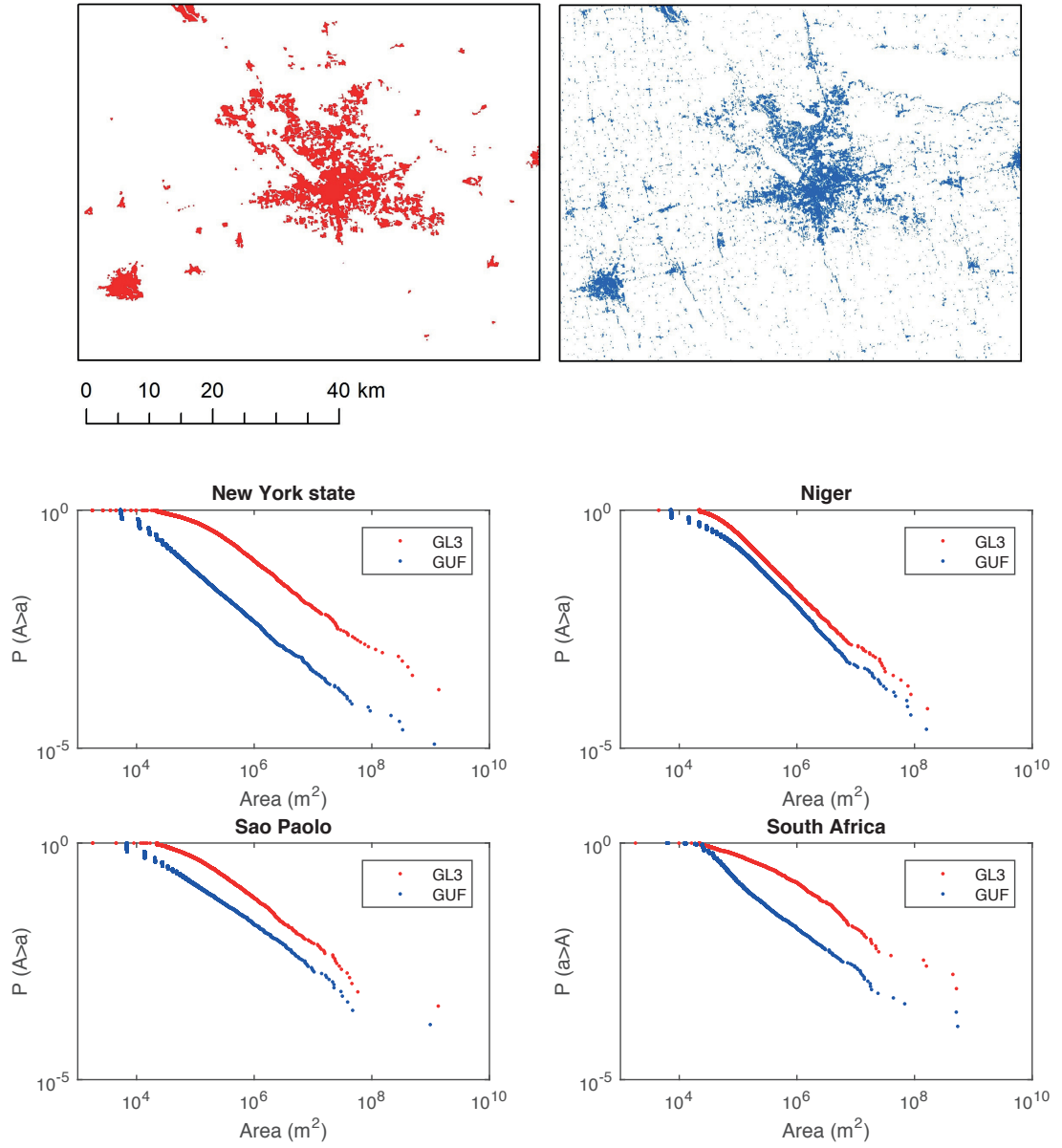


Figure 3.30: **Scaling for low and high resolution urban inventory.** The upper-left map shows the GL3 dataset, while the upper-right map shows the GUF data. Each plot shows the cumulative distribution of urban patches. The colours correspond to the map colours: red is GL3, and blue is GUF. The quantitative results are reported in Table3.31.

Chapter 3. Results

Table 3.2: The table reports basic quantities for the GL3 and the GUF data. $\sqrt{\min(A)}(m)$ is the square root of the area of the smaller urban patch, while $\sqrt{\max(A)}(m)$ is that of the bigger urban patch. $tot.A$ is the total area, and $\langle A \rangle$ is the average area. $\sqrt{X_{min}}$ is the lower cut-off of the fitted power law distribution expressed as $P(A) = a^{-\alpha}$.

Location	Data	$\sqrt{\min(A)}(m)$	$\sqrt{\max(A)}(m)$	$tot.A(km^2)$	$\langle A \rangle (Km^2)$	$\sqrt{X_{min}}(m)$	α
New York State	GUF	72	34019	5932	268	296	1.97
	GL3	316	37098	6376	1387	654	2.06
Niger	GUF	84	12629	25304	313	452	2.23
	GL3	316	12886	42107	707	326	2.25
Sao Paolo	GUF	82	31393	50989	534	320	1.84
	GL3	316	36711	38796	1390	450	1.91
South Africa	GUF	78	23283	47080	532	378	1.94
	GL3	316	22126	47557	1801	962	1.90

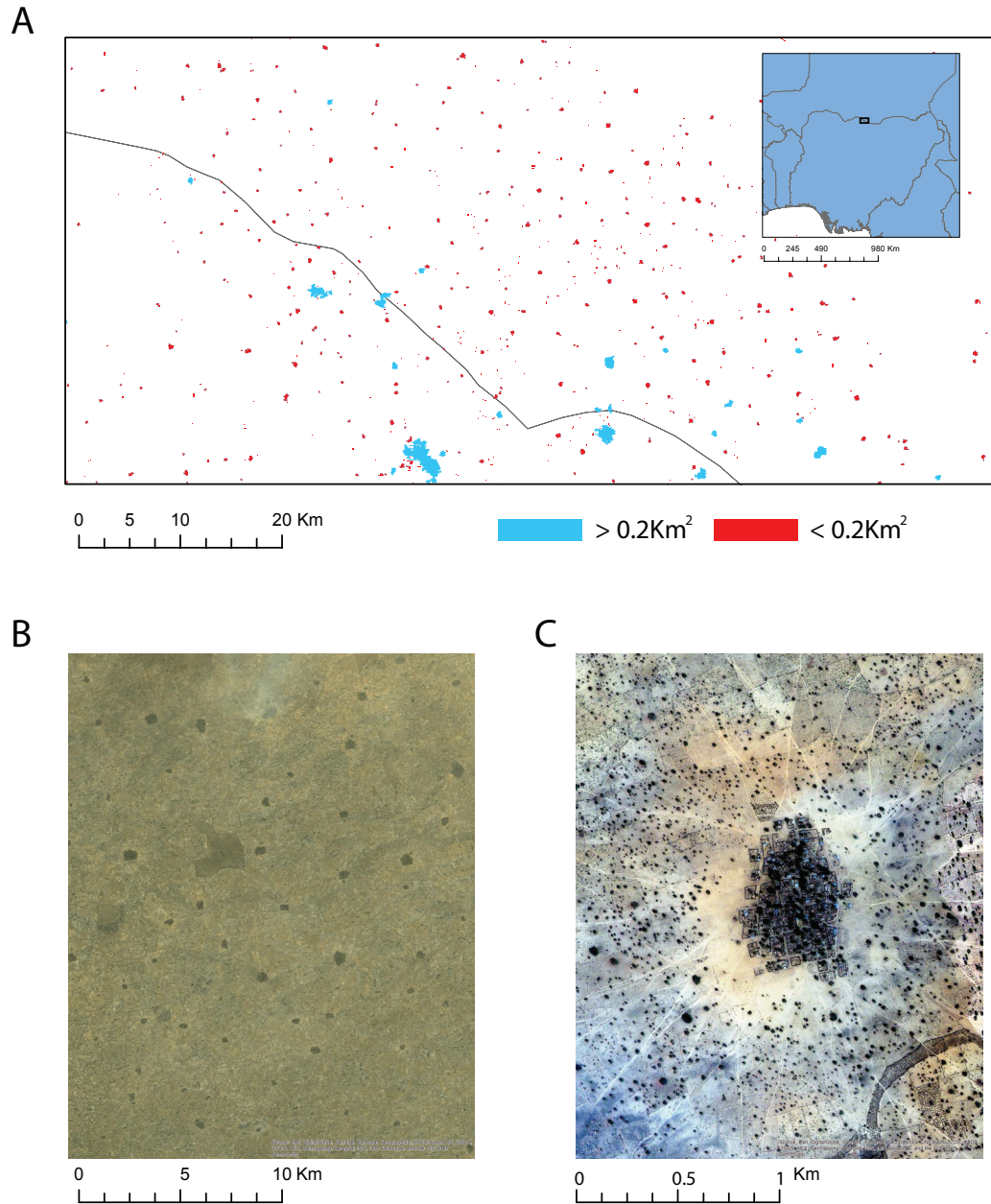


Figure 3.31: **Urban vs. rural patterns in Niger and Nigeria.**A) The map shows a portion of the Niger and Nigeria case studies. Red settlements have areas smaller than the lower cut-off (0.2km^2), as reported in Table 3.2. Blue settlements have areas of $> 0.2\text{km}^2$. B and C show only red settlements (smaller than the lower cut-off). In B and C, it is possible to observe the rural nature of the settlements and the absence of a paved road network.

3.6.3 Discussion

First, the absence of an upper bound of the cumulative distribution of all settlements on Earth indicates that cities, at a global scale, are far from their maximum size. In other words, the bigger cities on Earth can keep growing, and the limit of the urban systems might be the size of the system itself. In this sense, global urbanization can be considered to still be in a phase of growth and far from a steady condition. The absence of an upper cut-off also suggests that all cities on Earth might be a part of a unique process; therefore, global urbanization is a unique and single phenomena in which each city is somehow interdependent with all other cities. Regional scaling analysis, however, shows that in many highly urbanized regions, the distribution of settlement areas might not follow a power law. This is the case of the hyper-urbanized regions in China, which have probably reached a sort of saturation point. These preliminary results point to a deeper analysis aimed to fit different functional forms for different areas. The existence of different distributions suggests the possibility of classifying urban patterns in different typologies or classes of urbanization. Such an analysis needs to be done on full and more accurate data. The last analysis shows that the resolution of urban imagery is of great importance for an accurate analysis of the area distributions. A comparison between low and high resolution data shows that the lower bound of the power law can be used to discriminate between urban and rural settlement patterns. The detection of a precise minimum bound and the consequent objective distinction between urban and rural patterns can have great impact for two reasons: In the cases of New York state, Sao Paulo and South Africa, the lower cut-off bound is close to the minimum area. It is possible to claim that the urban/rural distinction does not exist, which could be why no one has been able to detect it so far. A second result is the existence of a clear minimum, as in the case of Niger and Nigeria, which implies the coexistence of two different settlements patterns that are in turn presumably generated by different dynamics. An objective distinction between rural and urban settlements may challenge and redraw the global urban transition figure shown in Figure1.

3.6.4 Perspective: A new inventory of global human settlements

All of the results show that urban scaling is still prone to further exploration and that accurate results depend on the accuracy of the data. These results stimulated a debate around the necessity of enlarging the GUF dataset to a global scale. Such data has been prepared together with DLR and is now available for further analysis. Figure 3.32 shows a preview of this data.



Figure 3.32: **A preview of GUF at global scale.** The image shows urbanized areas in Europe.

4 Conclusions and perspectives

4.1 Conclusions: Towards the development of a science for cities

Cities are the human object *par excellence*. Cities are the cause and effect of human cultural evolution and the scenario of the complex relationships between humans and nature. Cities are the final and perhaps the sole way in which human colonize Earth. Considering the global trend of urbanization, any debate on global issues, such as climate change, global energy demand or global social sustainability, is related to cities. It is evident, however, that traditional approaches of understanding cities are unfit for their complexity. The increasing necessity of a scientific understanding of cities [162, 163, 164] represents an avenue for building a new awareness of cities and urbanization. Understanding cities with paradigms from chemistry and physics is a relatively young tradition that, starting from the seminal and fundamental works of Jane Jacobs [64] and Michael Batty [165], developed in different scientific investigation branches. Today, problems related to cities and urbanization have been consistently labelled as problems of organized complexity [101].

Cities are complex in the way they have been built, resulting in a complex and self-organized spatial organization [165]. Cities are also complex in their internal dynamics. People, economy, trades, information, and all human actions and dynamics come from a highly decentralized and bottom-up planning system. However, cities are not closed systems; they interact with natural and non-urbanized environments. External interactions add a further layer of complexity to cities. Any investigation aiming to uncover the functional structure of towns must take into account such complexity and employ quantitative methodologies for analysis. The present study represents an effort in this direction; it proposes a series of analyses of town and urbanization patterns in which physics meets cities and where quantitative analyses corroborate qualitative information in an attempt to explain urban form and dynamics.

In Chapter 3.1, it has been demonstrated that the evolution of road networks is dominated by simple processes, which are mostly determined by spatial constraints. The processes of network growth are constant over a 200 year period and are thus independent of any transportation technology. The dynamics of road network growth (and not its rate of growth) cannot be related to car dependence or other transportation modalities. This might seem counter-intuitive, and such results must be tested on wider data. However, such results suggest that road networks are certainly not related to local or

economic factors, such as economy or land use. The universality of the road network structure was confirmed in Chapter 3.5, in which the network made by all roads on Earth were analysed. It was proven that its spatial organization is independent of land use and location. At least for road networks, the evidence proves that there are no specific relationships between the form of the network and its function. The global process of land fragmentation, as operated by road network evolution and at its stage of development, can be seen as a single, unique and universal phenomenon. This provides insights for a vision of urbanization as a process that must be taken as a global entity.

The study of the evolving street network also shows that the structural organization of roads, as described by centrality analysis, can trace back its historical evolution. Betweenness centrality is correlated with the age of the road. Centrality, and in particular betweenness centrality, has been used extensively in urban design practice. Its main use is to determine the impact of virtual flows of road pattern design. Betweenness, however, as shown in Chapter 3.4, does not mirror the multiplex nature of urban networks, as easily deducible underground systems play a crucial role in the spatial displacement or urban centralities. Empirical tests, as shown in Chapter 3.4, show that virtual flows are pushed out-of-town, suggesting that underground systems might be decentralized forces. However, they also increase the proximity of central places, simultaneously playing a centralized force. Change in the centrality displacement suggests that urban evolving forces, due to the interaction of layered transportation infrastructures, must be considered in any urban model that simulates urban evolution (for example, the model in Chapter 1.2).

The analysis of the global road network shows universal structural features among locations. Such universality could be explained in terms of planarity. However, as demonstrated in the comparative analysis of the simplicity index (Chapter 3.3), not all planar networks are similar. The urban network shows special geometric characteristics that differentiate them from other planar networks. We can deduce that road networks are all similar, but they represent a special sub-category of planar networks. In a way, it is not possible to claim that the spatial constraints imposed by planarity are the only factor that makes all road networks similar; other forces must have played a role in shaping them.

In Chapter 3.5 and 3.6, it has been shown that self-similarity is a constant and universal feature. This certainly implies that urbanization and fragmentation determinants operate with the same intensity at all scales. It might be possible, however, that such universal scaling is a result of the interaction between agglomeration forces (as in the case of urban settlements) together with fragmentation pressure (as in the case of road networks). Understanding this interaction might be a key analysis for producing a more reliable model of land use conversion.

Many problems still exist for a scientific understanding of cities and for providing a set of clear information that ultimately aims to improve cities. Regardless of specific scientific problems and further investigations, one of the bigger impediments to developing a science of cities is its relationship with urban planning and policy makers. There is a consistent and evident separation between contemporary urban planning practices, which rely mostly on qualitative considerations, and a more objective and scientific view of cities. A renovation of the entire urban planning discipline is necessary. Urban planners and policy makers must be aware of quantitative tools to understand cities. On the other hand, a positivist approach to towns, in which numbers lead the decision, can cause significant damage. The 1980s transportation engineering and modern urban planning, which spoiled many cities in Europe

4.2. Perspectives: The hypothesis of the global trade network as global urbanization determinant

and the US, leave a good lesson to remember.

Some general remarks might be done in term of statistical significance of power-law and self-similar probability distribution of urban settlements. It can be possible in fact, as proposed in Sec. 3.6, that beyond a main and coarse grained behaviour of functional forms, tiny and small differences between them can discriminate different patterns of urbanization. This is why, considering the accuracy of data, more statistical test are necessary to draw final conclusions on the use of urban scaling approach in the determination of urban patterns differences.

4.2 Perspectives: The hypothesis of the global trade network as global urbanization determinant

The thesis has demonstrated that, in terms of space and form, all cities are similar. In a way, cities in China and in Africa have very similar shapes. It seems that all city shapes are similar across scale and location. This hypothesis has been tested and confirmed using all cities on Earth. However, self-similarity across location does not explain the determinants of urbanization. A new research avenue aims to understand why cities continue to grow and why they are all, apparently, part of the same system, as shown in Chapter 3.6. An explanation of urbanization dynamics should be explored in economic terms and especially in the economic relationships between countries. It is an established fact that urbanization in developed countries was accompanied by economic growth and industrialization. This historic pattern generated the expectation of a virtuous circle between economic growth and urbanization regardless of local conditions [28, 29]. From classic urban economic theories [30, 31] to the more recent scaling approach to cities [59], the growth of urban populations was routinely used as a proxy of economic growth. This pattern has also been seen in developing countries such as China or India; however, it cannot be considered a universal pattern [32], and the deviations from it have not yet been fully explained. In fact, as noted by several studies [33, 16, 21, 35], the increasing urbanization rate in persistently poor and non-industrialized countries poses an important problem for urban economic theories. For example, why in Kinshasa and Dhaka, both with a population of more than 10 million, is the rate of economic growth almost stagnant? Why, given the same rate of urbanization, does Asia contain a number of highly explosive economies while sub-Saharan Africa has seen very little growth? In urban economics, there are several theories that aim to explain these uncommon patterns. Some say that rural poverty pushed people to cities [166], while others argue that in the last several decades, urban-biased policy led to over-urbanization [35]. The most intriguing explanation for these deviating patterns, first proposed by Krugman and Fujita [167], is rooted in the endogenous effect of the global trade network on local urbanization. We can call it the globalization hypothesis. The dominant idea behind the globalization hypothesis is that for open economies, domestic communities can trade with other communities, triggering an export substituting industrialization policy. Put simply, because commodities can flow much more freely than before, industrialization is not necessary to increase the net-product of a country because of the export-based economy. It has been speculated as the substitution industrialization-export of urbanization without industrialization or economic growth. The impact of this hypothesis on literature has been strong, and several studies have attempted to empirically test it. The most important contributions are those by Gollin and collaborators. They proved that agricultural sectors [21] and the export of natural resources [34] are positively associated

with urbanization in sub-Saharan Africa. However, despite this progress, there are some evident gaps. This hypothesis still needs to be proved. 1: In the analyses of trade network vs. urbanization, the definition of trade is always given by the mere summation of import or export commodities of a given type. This in turn precludes exploring other potentially important features of the global trade network as economic variables. This input comes directly from complex networks theory that extensively studied the structural properties of the global trade network. For example, it has been demonstrated that the global trade network shows a strongly hierarchical structure (given by a scale free distribution of node degree), dissortative behaviour (indicating the presence of strong hubs linked to weak nodes) and heterogeneity in the trade flow [168, 169, 170]. These observations point to an exploration of the effect of different network variables as determinants of urbanization. 2: Another possible gap, as argued by Montgomery [16], is that economic output has been taken as the driver for big cities or metropolitan areas that represents 40% of people. Medium or rural settlement growth have mostly not been considered. These observed gaps suggest further possible analyses to test the hypothesis of the global trade network as an urbanization determinant.

Bibliography

- [1] D. UN, "World population prospects: The 2012 revision," 2013.
- [2] "Un population division." <http://www.un.org/en/development/desa/population/>.
- [3] B. J. Berry, "Urbanization," in *Urban Ecology*, pp. 25–48, Springer, 2008.
- [4] K. C. Seto, R. Sánchez-Rodríguez, and M. Fragkias, "The new geography of contemporary urbanization and the environment," *Annual review of environment and resources*, vol. 35, pp. 167–194, 2010.
- [5] U. Habitat, "Cities in a globalizing world: global report on human settlements 2001," *London: Earthscan*, vol. 344, 2001.
- [6] U. Habitat, "The state of the world cities. globalization and urban culture," 2004.
- [7] U. Habitat, "State of world' s cities 2006/7," *New York: United Nations*, 2006.
- [8] E. L. Birch and S. M. Wachter, *Global urbanization*. University of Pennsylvania Press, 2011.
- [9] N. B. Grimm, S. H. Faeth, N. E. Golubiewski, C. L. Redman, J. Wu, X. Bai, and J. M. Briggs, "Global change and the ecology of cities," *science*, vol. 319, no. 5864, pp. 756–760, 2008.
- [10] D. Tilman, C. Balzer, J. Hill, and B. L. Befort, "Global food demand and the sustainable intensification of agriculture," *Proceedings of the National Academy of Sciences*, vol. 108, no. 50, pp. 20260–20264, 2011.
- [11] R. K. Kaufmann, K. C. Seto, A. Schneider, Z. Liu, L. Zhou, and W. Wang, "Climate response to rapid urban growth: evidence of a human-induced precipitation deficit," *Journal of Climate*, vol. 20, no. 10, pp. 2299–2306, 2007.
- [12] L. Zhou, R. E. Dickinson, Y. Tian, J. Fang, Q. Li, R. K. Kaufmann, C. J. Tucker, and R. B. Myneni, "Evidence for a significant urbanization effect on climate in china," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 26, pp. 9540–9544, 2004.
- [13] K. C. Seto and J. M. Shepherd, "Global urban land-use trends and climate impacts," *Current Opinion in Environmental Sustainability*, vol. 1, no. 1, pp. 89–95, 2009.
- [14] M. Moore, P. Gould, and B. S. Keary, "Global urbanization and impact on health," *International journal of hygiene and environmental health*, vol. 206, no. 4, pp. 269–278, 2003.

Bibliography

- [15] U. Habitat, "State of the world's cities 2006/7," *New York: United Nations*, 2006.
- [16] M. R. Montgomery, "The urban transformation of the developing world," *science*, vol. 319, no. 5864, pp. 761–764, 2008.
- [17] S. Angel, J. Parent, D. L. Civco, A. Blei, and D. Potere, "The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050," *Progress in Planning*, vol. 75, no. 2, pp. 53–107, 2011.
- [18] D. Potere and A. Schneider, "A critical look at representations of urban areas in global maps," *GeoJournal*, vol. 69, no. 1-2, pp. 55–80, 2007.
- [19] D. Potere, A. Schneider, S. Angel, and D. L. Civco, "Mapping urban areas on a global scale: which of the eight maps now available is more accurate?," *International Journal of Remote Sensing*, vol. 30, no. 24, pp. 6531–6558, 2009.
- [20] M. Batty, "The size, scale, and shape of cities," *science*, vol. 319, no. 5864, pp. 769–771, 2008.
- [21] D. Gollin, R. Jedwab, D. Vollrath, *et al.*, "Urbanization with and without industrialization," *Unpublished manuscript, Oxford University Department of International Development*, 2013.
- [22] W. Lutz and K. Samir, "Dimensions of global population projections: what do we know about future population trends and structures?," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 365, no. 1554, pp. 2779–2791, 2010.
- [23] K. C. Seto, M. Fragkias, B. Güneralp, and M. K. Reilly, "A meta-analysis of global urban land expansion," *PloS one*, vol. 6, no. 8, p. e23777, 2011.
- [24] K. C. Seto, B. Güneralp, and L. R. Hutyrá, "Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools," *Proceedings of the National Academy of Sciences*, vol. 109, no. 40, pp. 16083–16088, 2012.
- [25] M. Fragkias, B. Güneralp, K. C. Seto, and J. Goodness, "A synthesis of global urbanization projections," in *Urbanization, biodiversity and ecosystem services: Challenges and opportunities*, pp. 409–435, Springer, 2013.
- [26] J. Utzinger and J. Keiser, "Urbanization and tropical health then and now," *Annals of tropical medicine and parasitology*, vol. 100, no. 5-6, pp. 517–533, 2006.
- [27] A. J. Tatem, N. Campiz, P. W. Gething, R. W. Snow, and C. Linard, "The effects of spatial population dataset choice on estimates of population at risk of disease," *Population Health Metrics*, vol. 9, no. 1, p. 4, 2011.
- [28] M. Spence, P. C. Annez, and R. M. Buckley, *Urbanization and growth*. World Bank Publications, 2009.
- [29] G. Duranton, "Growing through cities in developing countries," *The World Bank Research Observer*, p. lku006, 2014.
- [30] J. Jacobs *et al.*, "The economy of cities.," *The economy of cities.*, 1970.

-
- [31] E. L. Glaeser, H. D. Kallal, J. A. Scheinkman, and A. Shleifer, "Growth in cities," tech. rep., National Bureau of Economic Research, 1991.
 - [32] M. Ravallion, S. Chen, and P. Sangraula, "New evidence on the urbanization of global poverty," *Population and Development Review*, vol. 33, no. 4, pp. 667–701, 2007.
 - [33] G. Petrakos, "Urbanization and international trade in developing countries," *World Development*, vol. 17, no. 8, pp. 1269–1277, 1989.
 - [34] D. Gollin, R. Jedwab, D. Vollrath, *et al.*, "Urbanization with and without structural transformation," *mimeograph*, George Washington University, 2013.
 - [35] E. L. Glaeser, "A world of cities: the causes and consequences of urbanization in poorer countries," *Journal of the European Economic Association*, vol. 12, no. 5, pp. 1154–1199, 2014.
 - [36] Q. Weng and D. A. Quattrochi, *Urban remote sensing*. CRC Press, 2006.
 - [37] E. Bartholomé and A. Belward, "Glc2000: a new approach to global land cover mapping from earth observation data," *International Journal of Remote Sensing*, vol. 26, no. 9, pp. 1959–1977, 2005.
 - [38] E. A. Bright, "Landscan global population 1998 database," *Oak Ridge National Laboratory, Oak Ridge, TN*, 1998.
 - [39] C. D. Elvidge, P. Cinzano, D. Pettit, J. Arvesen, P. Sutton, C. Small, R. Nemani, T. Longcore, C. Rich, J. Safran, *et al.*, "The nightsat mission concept," *International Journal of Remote Sensing*, vol. 28, no. 12, pp. 2645–2670, 2007.
 - [40] P. Defourny, C. Vancutsem, P. Bicheron, C. Brockmann, F. Nino, L. Schouten, M. Leroy, *et al.*, "Globcover: a 300 m global land cover product for 2005 using envisat meris time series," in *Proceedings of the ISPRS Commission VII mid-term symposium, Remote sensing: from pixels to processes*, pp. 8–11, Enschede, the Netherlands, 2006.
 - [41] K. Klein Goldewijk, A. Beusen, G. Van Drecht, and M. De Vos, "The hyde 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years," *Global Ecology and Biogeography*, vol. 20, no. 1, pp. 73–86, 2011.
 - [42] C. Small, "The color of cities: an overview of urban spectral diversity," *Global mapping of human settlements*, pp. 59–106, 2009.
 - [43] T. Esch, H. Taubenböck, A. Roth, W. Heldens, A. Felbier, M. Thiel, M. Schmidt, A. Müller, and S. Dech, "Tandem-x mission. new perspectives for the inventory and monitoring of global settlement patterns," *Journal of Applied Remote Sensing*, vol. 6, no. 1, pp. 061702–1, 2012.
 - [44] T. Esch, M. Marconcini, A. Felbier, A. Roth, W. Heldens, M. Huber, M. Schwinger, H. Taubenböck, A. Muller, and S. Dech, "Urban footprint processor. fully automated processing chain generating settlement masks from global data of the tandem-x mission," *Geoscience and Remote Sensing Letters, IEEE*, vol. 10, no. 6, pp. 1617–1621, 2013.
 - [45] L. M. A. M. A. Bettencourt, J. Lobo, and H. Youn, "The hypothesis of urban scaling: Formalization, implications and challenges," *SFI Working Paper*, 2013.

Bibliography

- [46] J. Brown, G. West, and B. Enquist, “Scaling in biology: Patterns and processes, causes and consequences, p. 1–24,” *Scaling in Biology. Oxford University Press, Oxford*, 2000.
- [47] X. Gabaix, “Zipf’s law for cities: an explanation,” *Quarterly journal of Economics*, pp. 739–767, 1999.
- [48] M. E. Newman, “Power laws, pareto distributions and zipf’s law,” *Contemporary physics*, vol. 46, no. 5, pp. 323–351, 2005.
- [49] G. K. Zipf, “Human behavior and the principle of least effort,,” 1949.
- [50] P. R. Krugman and P. Krugman, *The self-organizing economy*. Blackwell Oxford, 1996.
- [51] L. H. Dobkins, Y. M. Ioannides, *et al.*, “Dynamic evolution of the us city size distribution,” *The Economics of Cities*, pp. 217–260, 2000.
- [52] K. T. Rosen and M. Resnick, “The size distribution of cities: an examination of the pareto law and primacy,” *Journal of Urban Economics*, vol. 8, no. 2, pp. 165–186, 1980.
- [53] G. R. Carroll, “National city-size distributions what do we know after 67 years of research?,” *Progress in Human Geography*, vol. 6, no. 1, pp. 1–43, 1982.
- [54] B. J. Berry and A. Okulicz-Kozaryn, “The city size distribution debate: Resolution for us urban regions and megalopolitan areas,” *Cities*, vol. 29, pp. S17–S23, 2012.
- [55] P. Krugman, “Confronting the mystery of urban hierarchy,” *Journal of the Japanese and International economies*, vol. 10, no. 4, pp. 399–418, 1996.
- [56] M. Batty, “Rank clocks,” *Nature*, vol. 444, no. 7119, pp. 592–596, 2006.
- [57] D. Rybski, A. G. C. Ros, and J. P. Kropp, “Distance-weighted city growth,” *Physical Review E*, vol. 87, no. 4, p. 042114, 2013.
- [58] P. M. Romer, “Increasing returns and long-run growth,” *The journal of political economy*, pp. 1002–1037, 1986.
- [59] L. M. A. Bettencourt, J. Lobo, D. Helbing, C. Kuhnert, and G. B. West, “Growth, innovation, scaling, and pace of life in cities,” *Proceedings of National Academy of Sciences*, vol. 104, no. 17, pp. 7301–7306, 2007.
- [60] J. S. M. H. A. Oliveira Erneson A., Andrade Jr., “Large cities are less green,” *Sci. Rep.*, 2014.
- [61] E. Arcaute, E. Hatna, P. Ferguson, H. Youn, A. Johansson, and M. Batty, “City boundaries and the universality of scaling laws,” *arXiv preprint arXiv:1301.1674*, 2013.
- [62] L. M. Bettencourt, “The origins of scaling in cities,” *science*, vol. 340, no. 6139, pp. 1438–1441, 2013.
- [63] A. Clauset, C. R. Shalizi, and M. E. Newman, “Power-law distributions in empirical data,” *SIAM review*, vol. 51, no. 4, pp. 661–703, 2009.
- [64] J. Jacobs, “The life and death of great american cities,” 1961.

-
- [65] S. Lämmer, B. Gehlsen, and D. Helbing, "Scaling laws in the spatial structure of urban road networks," *Physica A Statistical Mechanics and its Applications*, vol. 363, pp. 89–95, 2006.
 - [66] S. Porta, P. Crucitti, and V. Latora, "The network analysis of urban streets: a primal approach," *ENVIRONMENT AND PLANNING B PLANNING AND DESIGN*, vol. 33, no. 5, p. 705, 2006.
 - [67] S. Porta, P. Crucitti, and V. Latora, "The network analysis of urban streets: A dual approach," *Physica A Statistical Mechanics and its Applications*, vol. 369, pp. 853–866, Sept. 2006.
 - [68] M. Barthélemy and A. Flammini, "Modeling urban street patterns," *Physical review letters*, vol. 100, no. 13, p. 138702, 2008.
 - [69] P. Crucitti, V. Latora, and S. Porta, "Centrality in networks of urban streets," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 16, no. 1, pp. –, 2006.
 - [70] M. Barthelemy, "Spatial networks," *Physics Reports*, vol. 499, no. 1–3, pp. 1 – 101, 2011.
 - [71] M. Barthelemy, P. Bordin, H. Berestycki, and M. Griboaudi, "Self-organization versus top-down planning in the evolution of a city," *Sci. Rep.*, vol. 3, p. 2153, 2012.
 - [72] L. da F Costa, B. A. N. Travençolo, M. P. Viana, and E. Strano, "On the efficiency of transportation systems in large cities," *Europhys. Lett.*, vol. 91, no. 1, p. 18003, 2010.
 - [73] S. Porta, V. Latora, and E. Strano, "Networks in urban design. six years of research in multiple centrality assessment," in *Network Science*, pp. 107–129, Springer London, 2010.
 - [74] Chan, S. H.Y., Donner, R. V., and Lämmer, S., "Urban road networks - spatial networks with universal geometric features?," *Eur. Phys. J. B*, 2011.
 - [75] T. Courtat, C. Gloaguen, and S. Douady, "Mathematics and morphogenesis of cities: A geometrical approach," *Physical Review E*, vol. 83, no. 3, p. 036106, 2011.
 - [76] J. Buhl, J. Gautrais, N. Reeves, R. V. Sole, P. K. S. Valverde, and G. Theraulaz, "Topological patterns in street networks of self organized urban settlements," *Eur. Phys. J. B*, vol. 49, p. 513, 2006.
 - [77] A. P. Masucci, E. Arcaute, E. Hatna, K. Stanilov, and M. Batty, "On the problem of boundaries and scaling for urban street networks," *Journal of The Royal Society Interface*, vol. 12, no. 111, p. 20150763, 2015.
 - [78] A. Cardillo, S. Scellato, V. Latora, and S. Porta, "Structural properties of planar graphs of urban street patterns," , vol. 73, no. 6, pp. 066107–+, 2006.
 - [79] A. P. Masucci, D. Smith, A. Crooks, and M. Batty, "Random planar graphs and the london street network," *The European Physical Journal B*, vol. 71, no. 2, pp. 259–271, 2009.
 - [80] E. Strano, M. Viana, A. Cardillo, L. D. F Costa, S. Porta, and V. Latora, "Urban street networks, a comparative analysis of ten european cities.," *Environ. Plann. B*, vol. 40, no. 6, pp. 1071–1086, 2013.
 - [81] S. Wasserman and K. Faust, *Social Network Analysis*. Cambridge: Cambridge University Press, 1994.

Bibliography

- [82] C. L. Freeman *Sociometry*, vol. 40, no. 35, 1977.
- [83] C. L. Freeman *Social Networks*, vol. 1, no. 215, 1979.
- [84] M. M. V. Latora *Phys. Rev. Lett.*, vol. 87, no. 198701, 2001.
- [85] S. B. V. L. Y. M. M. Chavez and D.-U. Hwang, “Complex networks: Structure and dynamics,” *Phys. Rep.*, vol. 424, no. 175, 2006.
- [86] J. Scott, *Social Network Analysis: A Handbook*. London (UK): Sage, 2nd ed., 2000.
- [87] G. Sabidussi, “The centrality index of a graph,” *Psychometrika*, no. 31, pp. 581–603, 1966.
- [88] S. Porta, V. Latora, F. Wang, E. Strano, A. Cardillo, S. Scellato, V. Iacoviello, and R. Messori, “Street centrality and densities of retail and services in Bologna, Italy,” *Environ. Plann. B*, vol. 36, no. 3, pp. 450–465, 2009.
- [89] “Open street map.” <http://www.openstreetmap.org>.
- [90] “Geoopen.” <https://www.sharegeo.ac.uk/>.
- [91] C. C. University, “Global roads open access data set, version 1 (groadsv1),” 2013.
- [92] “delorme.” <http://www.delorme.com/digitalmapdata/world.htm>.
- [93] “Urbanatlas.” <http://www.eea.europa.eu/legal/copyrigh>.
- [94] J. Chen, J. Chen, A. Liao, X. Cao, L. Chen, X. Chen, C. He, G. Han, S. Peng, M. Lu, *et al.*, “Global land cover mapping at 30m resolution: A pok-based operational approach,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 103, pp. 7–27, 2015.
- [95] S. Fritz, L. See, I. McCallum, L. You, A. Bun, E. Moltchanova, M. Duerauer, F. Albrecht, C. Schill, C. Perger, *et al.*, “Mapping global cropland and field size,” *Global change biology*, vol. 21, no. 5, pp. 1980–1992, 2015.
- [96] “geowiki.” <http://www.geo-wiki.org/>.
- [97] B. L. Turner, E. F. Lambin, and A. Reenberg, “The emergence of land change science for global environmental change and sustainability,” *Proceedings of the National Academy of Sciences*, vol. 104, no. 52, pp. 20666–20671, 2007.
- [98] C. E. Ramalho and R. J. Hobbs, “Time for a change: dynamic urban ecology,” *Trends in ecology & evolution*, vol. 27, no. 3, pp. 179–188, 2012.
- [99] A. G. Wilson, *Entropy in urban and regional modelling*. Pion Ltd, 1970.
- [100] M. Batty and P. A. Longley, *Fractal cities: a geometry of form and function*. Academic Press, 1994.
- [101] M. Batty, *Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals*. The MIT press, 2007.
- [102] H. A. Makse, S. Havlin, and H. E. Stanley, “Modelling urban growth patterns,” *Nature*, vol. 377, no. 6550, pp. 608–612, 1995.

-
- [103] N. A. Salingaros, *Principles of urban structure*, vol. 4. Techne Press, 2005.
 - [104] H. Taubenböck, T. Esch, A. Felbier, M. Wiesner, A. Roth, and S. Dech, "Monitoring urbanization in mega cities from space," *Remote sensing of Environment*, vol. 117, pp. 162–176, 2012.
 - [105] S. Marshall, *Streets and patterns*. Routledge, 2004.
 - [106] M. Southworth and E. Ben-Joseph, *Streets and the Shaping of Towns and Cities*. Island Press, 2003.
 - [107] E. Turri, *La megalopoli padana*. Marsilio, 2000.
 - [108] M. Barthélemy and A. Flammini, "Modeling urban street patterns," *Phys. Rev. Lett.*, vol. 100, no. 13, p. 138702, 2008.
 - [109] M. Barthélemy, "Spatial networks," *Physics Reports*, vol. 499, no. 1, pp. 1–101, 2011.
 - [110] P. Haggett and R. J. Chorley, *Network analysis in geography*, vol. 67. Edward Arnold London, 1969.
 - [111] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, pp. 35–41, 1977.
 - [112] P. Crucitti, V. Latora, and S. Porta, "Centrality measures in spatial networks of urban streets," , vol. 73, pp. 036125–+, Mar. 2006.
 - [113] M. C. Comerio, "Built for change: Neighborhood architecture in san francisco," *Journal of Architectural Education*, vol. 43, no. 1, pp. 51–53, 1989.
 - [114] J. Clark and D. A. Holton, *A first look at graph theory*, vol. 1. World Scientific, 1991.
 - [115] W. T. Tutte, "A census of planar maps," *Canad. J. Math*, vol. 15, no. 2, pp. 249–271, 1963.
 - [116] Y. Mileyko, H. Edelsbrunner, C. A. Price, and J. S. Weitz, "Hierarchical ordering of reticular networks," *PloS one*, vol. 7, no. 6, p. e36715, 2012.
 - [117] E. Katifori and M. O. Magnasco, "Quantifying loopy network architectures," *PloS one*, vol. 7, no. 6, p. e37994, 2012.
 - [118] M. Rosvall, A. Trusina, P. Minnhagen, and K. Sneppen, "Networks and cities: An information perspective," *Physical Review Letters*, vol. 94, no. 2, p. 028701, 2005.
 - [119] R. Dorfman, "A formula for the gini coefficient," *The Review of Economics and Statistics*, pp. 146–149, 1979.
 - [120] E. Strano, V. Nicosia, V. Latora, S. Porta, and M. Barthélemy, "Elementary processes governing the evolution of road networks," *Sci. Rep.*, vol. 2, 2012.
 - [121] A. Tero, S. Takagi, T. Saigusa, K. Ito, D. P. Bebbber, M. D. Fricker, K. Yumiki, R. Kobayashi, and T. Nakagaki, "Rules for biologically inspired adaptive network design," *Science*, vol. 327, no. 5964, pp. 439–442, 2010.
 - [122] V. Latora and M. Marchiori, "Is the Boston subway a small-world network?," *Physica A: Statistical Mechanics and its Applications*, vol. 314, no. 1–4, pp. 109 – 113, 2002.

Bibliography

- [123] S. Derrible and C. Kennedy, “Network analysis of world subway systems using updated graph theory,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2112, no. 1, pp. 17–25, 2009.
- [124] S. Derrible and C. Kennedy, “Characterizing metro networks: state, form, and structure,” *Transportation*, vol. 37, no. 2, pp. 275–297, 2010.
- [125] C. Roth, S. M. Kang, M. Batty, and M. Barthelemy, “A long-time limit for world subway networks,” *J. R. Soc. Interface*, 2012.
- [126] S. Derrible and C. Kennedy, “The complexity and robustness of metro networks,” *Physica A: Statistical Mechanics and its Applications*, vol. 389, no. 17, pp. 3678–3691, 2010.
- [127] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, “Catastrophic cascade of failures in interdependent networks,” *Nature*, vol. 464, pp. 1025–1028, 2010.
- [128] M. Kivelä, A. Arenas, M. Barthélemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, “Multilayer networks,” *arXiv:1309.7233*, 2013.
- [129] S. Shai and S. Dobson, “Coupled adaptive complex networks,” *Phys. Rev. E*, vol. 87, no. 4, p. 042812, 2013.
- [130] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M. A. Porter, S. Gómez, and A. Arenas, “Mathematical formulation of multilayer networks,” *Phys. Rev. X*, vol. 3, no. 4, p. 041022, 2013.
- [131] S. Boccaletti, G. Bianconi, R. Criado, C. del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, and M. Zanin, “The structure and dynamics of multilayer networks,” *Physics Reports*, 2014.
- [132] S. Gómez, A. Díaz-Guilera, J. Gómez-Gardeñes, C. J. Pérez-Vicente, Y. Moreno, and A. Arenas, “Diffusion dynamics on multiplex networks,” *Phys. Rev. Lett.*, vol. 110, no. 2, p. 028701, 2013.
- [133] M. De Domenico, A. Solé-Ribalta, S. Gómez, and A. Arenas, “Navigability of interconnected networks under random failures,” *Proc. Natl. Acad. Sci. U.S.A.*, vol. 111, no. 23, pp. 8351–8356, 2014.
- [134] R. G. Morris and M. Barthelemy, “Transport on coupled spatial networks,” *Phys. Rev. Lett.*, vol. 109, no. 12, p. 128703, 2012.
- [135] R. Gallotti and M. Barthelemy, “in press,” *Sci. Rep.*, 2014.
- [136] R. Z. Farahani, E. Miandoabchi, W. Szeto, and H. Rashidi, “A review of urban transportation network design problems,” *European Journal of Operational Research*, vol. 229, no. 2, pp. 281 – 302, 2013.
- [137] P. Wang, T. Hunter, A. M. Bayen, K. Schechtner, and M. C. González, “Understanding road usage patterns in urban areas,” *Scientific reports*, vol. 2, 2012.
- [138] D. Levinson, “Density and dispersion: the co-development of land use and rail in london,” *Journal of Economic Geography*, vol. 8, pp. 55–77, 2008.

-
- [139] D. King, "Developing densely: Estimating the effect of subway growth on new york city land uses," *Journal of Transport and Land Use*, vol. 4, no. 2, 2011.
 - [140] M. Batty, P. Ferguson, *et al.*, "Defining city size," *Environment and Planning B: Planning and Design*, vol. 38, no. 5, pp. 753–756, 2011.
 - [141] U. C. Bureau, "Metropolitan and micropolitan statistical areas," 2008.
 - [142] "Eurostat." <http://ec.europa.eu/eurostat>.
 - [143] J. Svejnar, "Transition economies: Performance and challenges," *Journal of Economic Perspectives*, pp. 3–28, 2002.
 - [144] B. Heyns, "Emerging inequalities in central and eastern europe," *Annual review of sociology*, pp. 163–197, 2005.
 - [145] S. G. Perz, "Sustainable development: The promise and perils of roads," *Nature*, vol. 513, no. 7517, pp. 178–179, 2014.
 - [146] N. M. Haddad, "Corridors for people, corridors for nature," *Science*, vol. 350, no. 6265, pp. 1166–1167, 2015.
 - [147] D. F. Bryceson, A. Bradbury, and T. Bradbury, "Roads to poverty reduction? exploring rural roads' impact on mobility in africa and asia," *Development Policy Review*, vol. 26, no. 4, pp. 459–482, 2008.
 - [148] E. Strano, V. Nicosia, V. Latora, S. Porta, and M. Barthélemy, "Elementary processes governing the evolution of road networks," *Sci. Rep.*, vol. 2, 2012.
 - [149] W. F. Laurance, G. R. Clements, S. Sloan, C. S. OConnell, N. D. Mueller, M. Goosem, O. Venter, D. P. Edwards, B. Phalan, A. Balmford, *et al.*, "A global strategy for road building," *Nature*, 2014.
 - [150] J. Dulac, "Global land transport infrastructure requirements. estimating road and railway infrastructure capacity and costs to 2050," *International Energy Agency*, 2013.
 - [151] W. F. Laurance, "Bad roads, good roads," *Handbook of Road Ecology*, p. 449, 2015.
 - [152] P. Hagget and R. Chorley, "Network analysis in geography, vol. 348," *Edward Arnold, London, UK*, 1972.
 - [153] C. D. Elvidge, J. Safran, B. Tuttle, P. Sutton, P. Cinzano, D. Pettit, J. Arvesen, and C. Small, "Potential for global mapping of development via a nightsat mission," *GeoJournal*, vol. 69, no. 1-2, pp. 45–53, 2007.
 - [154] G. F. Jenks, "The data model concept in statistical mapping," *International yearbook of cartography*, vol. 7, no. 1, pp. 186–190, 1967.
 - [155] M. E. Fisher and M. N. Barber, "Scaling theory for finite-size effects in the critical region," *Physical Review Letters*, vol. 28, no. 23, p. 1516, 1972.
 - [156] J. R. Banavar, J. L. Green, J. Harte, and A. Maritan, "Finite size scaling in ecology," *Physical Review Letters*, vol. 83, no. 20, p. 4212, 1999.

Bibliography

- [157] J. R. Banavar, A. Maritan, and A. Rinaldo, “Size and form in efficient transportation networks,” *Nature*, vol. 399, p. 130, 1999.
- [158] A. Giometto, F. Altermatt, F. Carrara, A. Maritan, and A. Rinaldo, “Scaling body size fluctuations,” *Proceedings of the National Academy of Sciences*, vol. 110, no. 12, pp. 4646–4650, 2013.
- [159] S. M. Bhattacharjee and F. Seno, “A measure of data collapse for scaling,” *Journal of Physics A: Mathematical and General*, vol. 34, no. 33, p. 6375, 2001.
- [160] W. F. Laurance, L. V. Ferreira, J. M. Rankin-de Merona, and S. G. Laurance, “Rain forest fragmentation and the dynamics of amazonian tree communities,” *Ecology*, vol. 79, no. 6, pp. 2032–2040, 1998.
- [161] W. F. Laurance, P. Delamônica, S. G. Laurance, H. L. Vasconcelos, and T. E. Lovejoy, “Conservation: rainforest fragmentation kills big trees,” *Nature*, vol. 404, no. 6780, pp. 836–836, 2000.
- [162] M. Batty, “Building a science of cities,” *Cities*, vol. 29, pp. S9–S16, 2012.
- [163] M. Batty, *The new science of cities*. Mit Press, 2013.
- [164] W. Solecki, K. C. Seto, and P. J. Marcotullio, “It’s time for an urbanization science,” *Environment: Science and Policy for Sustainable Development*, vol. 55, no. 1, pp. 12–17, 2013.
- [165] M. Batty and P. A. Longley, *Fractal cities: a geometry of form and function*. Academic Press, 1994.
- [166] A. F. Ades and E. L. Glaeser, “Trade and circuses: explaining urban giants,” tech. rep., National Bureau of Economic Research, 1994.
- [167] M. Fujita, P. R. Krugman, and A. Venables, *The spatial economy: Cities, regions, and international trade*. MIT press, 2001.
- [168] D. Garlaschelli, T. Di Matteo, T. Aste, G. Caldarelli, and M. I. Loffredo, “Interplay between topology and dynamics in the world trade web,” *The European Physical Journal B*, vol. 57, no. 2, pp. 159–164, 2007.
- [169] G. Fagiolo, J. Reyes, and S. Schiavo, “The evolution of the world trade web: a weighted-network analysis,” *Journal of Evolutionary Economics*, vol. 20, no. 4, pp. 479–514, 2010.
- [170] L. De Benedictis and L. Tajoli, “The world trade network,” *The World Economy*, vol. 34, no. 8, pp. 1417–1454, 2011.

Curriculum vitae Emanuele Strano

March, 2016

Nationality: Italian
Data of birth: 10 July 1981
email: emanuele.strano@epfl.ch
Address : EPFL ENAC IIE LASIG, Bât. GC, St. 18, CH-1015 Lausanne,



SCIENTIFIC AEREAS

Global urbanization studies, urban science, transportation networks, urban planning and urban design.
Complex spatial networks, urban remote sensing, spatial statistic.

CURRENT POSITION

PhD Candidate

Laboratory of geographic Information Systems (LASIG),
École polytechnique Fédérale de Lausanne (EPFL),
Lausanne, CH.
Supervisors: Prof. Francois Golay (LASIG, EPFL, Lausanne, CH), Prof. Andrea Rinaldo (ECHO, EPFL, Lausanne, CH).

PAST POSITIONS

10-2009 – 10-2011

Research assistant

University of Strathclyde, Department of Architecture. Glasgow, UK.

12-2010 – 10-2011

Research associate

Oxford Brookes University, Department Of Planning. Oxford, UK.

EDUCATION

07-2010 – 07-2011

MRes in Architecture

University of Strathclyde, Glasgow, UK.
Title of the thesis: *Complex Networks and Urban Form*.

10-2008 – 02-2009

Internship

United Nation, Economic Commission For Europe, Housing and Land Management Unit. Geneva, CH.

07-2005 – 07- 2010

BSc in Architecture

Polytechnic of Milan, Italy.

THEACHING

2013-2014

Assistant at the course, “Geographic information system for decision making” .
EPFL, Master in Environmental Engineering.

2013-2015

Teaching at the course, “Geo-computation”
EPFL, Master in Environmental Engineering.

2014-2015

Assistant at the course, “Venice future cities” .
EPFL, ENAC.

ACADEMIC SERVICES

2015

Main organizer

Urban Science Days. Series of lectures during spring and fall semester 2015, funded by ENAC, Seed money grant.

29/30-08-2011

Main organizer and scientific director.

Workshop “*Evolution of Complex Transportation Networks*”, 29-30 August 2012 Strathclyde University, Glasgow. Funded by The University of Strathclyde, Science and Engineering Department.

Referee for:

Nature Scientific Report, Environment and Planning B, Phys. A, European Physical Journal Special Topic, Geographical Analysis, Spatial Statistics, International Journal of Geographic Information Science.

GRANTS

2016

Early Post-Doc Mobility Grant, Swiss National Foundation

Project title: Revisiting the global urbanization extent and the hypothesis of global trade network as its economic determinants.

Amount: 87000 CHF

Period: June- 2016, November 2016

Hosting institute: HuMnet Laboratory, Civil and Environmental Engineering Department, Massachusetts Institute of Technology (MIT), Boston, US

2015

Seed money on innovative research on cities. ENAC, EPFL.

Project title: Strano, E. Golay, F. Bertuzzo, E. Viganò, P. Rinaldo, A. Exploration in Urban Scaling

Amount: 60000 CHF

Period: Sept- 2014, Sept 2015

2014

Mobility Grant

Visiting scholar at Arizona State University.

Amount: 1000 CHF

Period: August 2014, November 2015.

GENERAL SKILLS

Languages

Italian: native

English; excellent

PUBLICATIONS

Published

- **Strano E**, Sood V, (2016) Urban scaling and the European economic transition. PlosOne (in print).
- Fiasconaro A, **Strano E**, Nicosia V, Porta S, Latora V (2016) *Spatio-temporal change of micro economic activities in Rome reveals patterns of mixed-use urban evolution*. PlosOne (in print).
- **Strano E**, Shay S, Dobson S, Barthelemy M, (2015) *Multiplex networks in metropolitan areas: generic features and local effects*. Royal Society Interface, 12 (111), 20150651.
- Viana M, **Strano E**, Bordin P, Barthelemy M, (2013) *The simplicity of planar graph*. Nature Scientific Report 4.
- Austwick M.Z, O’Brien O, **Strano E**, Viana M, (2013) *The structure of spatial networks and communities in bicycle sharing systems*, PloSone 89.

- Porta S, **Strano E**, et al. (2013) *Multiple centrality assessment*. Urban Design (125). pp. 12-14.
- **Strano E**, Viana M, Cardillo A, Porta S, Da Costa L, Latora V, (2013) *Urban street networks, a comparative analysis of ten European cities*. Environment and Planning B.
- **Strano E**, Nicosia V, Porta S, Latora V, Barthélemy M, (2012) *Elementary processes governing the evolution of roads networks*. Nature Scientific Report 2.
- **Strano E**, Adamatzky A, Jones J, (2012) Vie Physarale: Evaluation of Roman Roads with Slime Mould. International Journal of Nano Molecular Computing.
- Porta S, Latora V, Wang F, Rueda S, **Strano E**, Scellato S, Cardillo A, Belli E. (2011) *Street Centrality and Location of Economic Activities in Barcelona*. Urban Studies 10-1177.
- Costa L, Travencolo B, Viana M, **Strano E**, (2010) *On the Efficiency of Underground Systems in Large Cities*. Europhysics Letters 91-18003.
- Porta S, **Strano E**, Latora V, Wang F, Cardillo A, Scellato S, Iacoviello V, Messori R, (2009) *Street centrality and densities of retails and services in Bologna, Italy*. Environment and Planning B, 36 450- – 465.

Books

- Pooley C.G, Jones T, Tight M, Horton D, Scheldeman G, Mullen C, Jopson A, **Strano E**. (2013) Promoting walking and cycling. New perspectives on sustainable travel (Policy Press, Bristol, UK).

Books chapters

- Jones T, Horton D, Mullen C, Pooley C, Scheldeman G, **Strano E**. and Tight M. The Role of Street Network Connectivity and Access to Everyday Facilities in Shaping Everyday Walking and Cycling in English Cities.
In Gerike, R., Hulsmann, F. and Roller, K. (2013) *Strategies for Sustainable Mobilities: Opportunities and Challenges*. Farnham, Surrey: Ashgate.
- Porta S, Latora V, **Strano E**, (2010) *Networks in Urban Design. Six Years of Research in Multiple Centrality Assessment*.
In Estrada E, Fox M, Higham D, Oppo G, *Complex Network across the Natural and Technological Sciences*. Springer, London.

Conference proceedings

- **Strano E**, Shay S, *Structural analysis of the global roads network and the its impact on population, economy and urbanization*. VIII International PhD Seminar Urbanism & Urbanization | Symposium Latsis 2015, EPFL.
- Dibble J, Prelorndjos A, Romice O, Zanella M, **Strano E**, Pagel M, Porta S. *Urban Morphometrics: Towards a Science of Urban Evolution*. ISUF (International Society of Urban Form) conference Rome 22-26 Sept 2015.
- Produit T, Lachance N, **Strano E**, Porta S, Joost S, *A network based kernel density estimator applied to Barcelona economic activities*. ICCSA 2010, International Conference on Computational Science and Its Applications. Kyushu Sangyo University, Fukuoka, Japan. Volume I, Springer, London.

Edited volumes

- **Strano E**, Zanin M, Lillo F, Estrada E, (2012) *Spatially embedded socio-technical complex Networks*. European Physic Journal Special Topics.

Editorials

- **Strano E**, Zanin M, Lillo F, Estrada E, (2012) *Spatially embedded socio-technical complex networks*. European Physic Journal Special Topics.

Reports

- Tatem A, Sorichetta A, Bird R, Tejedor N, **Strano E**, Viana M, (2015) *Defining Regional Blocks for Malaria Elimination*. Mid-project progress report for grant OPP1115575. Bill and Melinda Gates Foundation.

International Conferences

- **Strano E**, Bertuzzo E, Shay S, Giometto A, Mucha P, Rinaldo A, *The structure and geography of the global road network, universal features and local urbanization processes*. Urban Transition Global Summit 2016m Shanghai, China.
- Esch T, **Strano E**, Heldens W, Hirner A, Keil M, Marconcini M, Roth A, Zeidler J, *The Global Urban Footprint Initiative - Pushing the Limits of Mapping Global Human Settlements from Space*. Urban Transition Global Summit 2016m Shanghai, China.
- **E. Strano**, Shay S, Dobson S, Barthelemy M, *Interdependent spatial networks in metropolitan areas*. Satellite workshop Urban Systems and Network, NetSci 2014, Berkeley, California.
- **E. Strano**, M. Austwick, M. Viana and O. O'Brien. *The effect of urban form on urban mobility*. Workshop on Green and Integrated Transportation Networks, ECCS2012, Brussels.
- S. Porta, R. Adam, O. Romice, **E. Strano**, A. Venerandi, P. Russell, *Multiple Centrality Assessment: understanding the evolution of cities and applying knowledge to practice in professional urban design*. Applied Urban Modelling: Assessing Pathways Towards Energy Efficient and Climate-Wise City Region, AUM2012,
- **E. Strano**. *Evolution of street network, a 200 years old empirical case study*. AGILE 2012 Workshop on Complexity Modeling for Urban Structure and Dynamics, Avignon (France) .
- M. Austwick, **E. Strano**, M. Viana and O. O'Brien. *Network approaches to journeys on the London bicycling sharing system*. AGILE 2012 Workshop on Complexity modelling for urban structure and dynamics, Avignon (France).
- **Strano E**, Jones T, Prouditi T, *The relationship between urban street network configuration, land use and short urban travel behaviour*. ECCS2011, Vienna. Satellite workshop, Complexity and the Future of Transportation Systems
- Jones T, **Strano E**, Jopson A, *Understanding the Relationship Between Urban Form and Walking and Cycling*. Royal Geographical Society Annual International Conference 2011, London, UK.
- Jones T, **Strano E**, *The contribution of urban form in encouraging everyday walking and cycling in English cities*. mobil.TUM 2011 conference. Making Sustainable Motilities - Interdisciplinary Perspectives, Munich, Germany.
- **Strano E**, Jones J, Adamatzky A, *Omnia Phisara Romam Ducunt, Physarum Policephalum Vs Vie Consolari*. Artificial Life XII 2010, 12th International Conference on the Synthesis and Simulation of Living Systems, Odense, Denmark.

